A Study of Prospective Secondary Mathematics Teachers' Evolving Understanding of Reasoning-and-Proving

by

Justin D. Boyle

B.S. Mathematics, University of Pittsburgh, 1997

MEd Mathematics Education, University of Pittsburgh, 2008

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

2012

UNIVERSITY OF PITTSBURGH

School of Education

This dissertation was presented

by

Justin D. Boyle

It was defended on

July 20, 2012

and approved by

Dr. James Greeno, Emeritus Professor, Stanford University

Dr. Ellen Ansell, Associate Professor, Instruction and Learning

Dr. Charles Munter, Assistant Professor, Instruction and Learning

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

A Study of Prospective Secondary Mathematics Teachers' Evolving Understanding of Reasoning-and-Proving

Justin D. Boyle, EdD

University of Pittsburgh, 2012

Proof is a foundational mathematical activity that has been underrepresented in school mathematics. The recently adopted Common Core State Standards in Mathematics includes eight process standards, several of which promote the inclusion of reasoning and proof across all grades, courses, and students. If students are to reach the expectations recommended by mathematics researchers and explicitly identified in the Common Core State Standards, then students will need opportunities to construct and validate proof arguments. However, secondary students find it challenging to validate arguments and produce proofs and do not know what a mathematical proof is. Furthermore, those preparing to be secondary mathematics teachers in undergraduate mathematics courses are unable to construct proofs on a consistent basis, and practicing secondary teachers possess a limited conception of proof.

A six-week graduate-level course was taught with the purpose of increasing practicing mathematics teachers' knowledge, expanding their conceptions of reasoning and proof, and preparing them to create similar experiences for their students. Research was conducted on the course to study the participants' evolving understanding of reasoning-and-proving. The results suggest that: 1) the course was successful at expanding the participants conception of proof; 2) the prospective teachers encountered five challenges when asked to write proofs that are at the secondary mathematics level; 3) specific types of arguments were challenging for participants to classify as proofs or non-proofs; and 4) even though the participants were skillful in selecting

high-level tasks that they could modify to include reasoning-and-proving opportunities, more work is needed to integrate such task across any secondary curricula.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS XII

1.0 CHAPTER 1: STATEMENT OF THE PROBLEM 1
1.1 INTRODUCTION 1
1.2 BACKGROUND 4
1.2.1 Policymakers Support Reasoning and Sense Making in Schools 5
1.2.2 What is proof? And what are the recommendations for it in secondary mathematics? .. 8
1.2.3 Students' Ability to Prove 9
1.2.4 Treatment of Proof in School 12
1.2.5 Teachers' Ability to Write Proof 14
1.2.6 Teachers and Students Harbor Deep Misconceptions of Proof 16
1.2.7 Conclusion 18
1.3 POLICY, PROFESSIONAL DEVELOPMENT, AND THE TEACHER COURSE 19
1.4 PURPOSE OF THIS STUDY 20
1.5 RESEARCH QUESTIONS 20
1.6 SIGNIFICANCE 21
1.7 LIMITATIONS 22
1.8 OVERVIEW 23

2.0 CHAPTER 2: REVIEW OF LITERATURE 24
2.1 LEARNING AND TEACHING PROOF 25
2.1.1 Secondary Students' Experiences and Understanding of Proof 26
2.1.1.1 Outside Classroom Assessment: Are Examples Convincing to Students? 27
2.1.1.2 Secondary classroom handling of proof discourse 32
2.1.2 College Students' and Teachers' Experiences with Learning to Prove 36
2.2 REASONING-AND-PROVING TASKS 42
2.3 THEORETICAL PERSPECTIVE FOR PREPARING TEACHERS TO IMPLEMENT R&P TASKS 51
2.4 THREE EXAMPLES OF PRODUCTIVE PROFESSIONAL DEVELOPMENT THAT EXPANDED TEACHER KNOWLEDGE 54
2.4.1 Educational Leaders in Mathematics (ELM) 57
2.4.2 Cognitively Guided Instruction (CGI) 61
2.4.3 Enhancing Secondary Mathematics Teacher Preparation (ESP) 66
2.5 SITUATING THE RESEARCH IN THIS CURRENT STUDY 71
2.5.1 Features of Productive Professional Development Programs 71
2.5.1.1 Learning theory 73
2.5.1.2 Explicit use of organizing frameworks 75
2.5.2 Grounding study in a theoretical model 77

3.0 CHAPTER 3: METHODOLOGY 80
3.1 CONTEXT OF R&P COURSE 81
3.1.1 The reasoning-and-proving course 81

3.1.1.1 Course activities ... 82
3.1.1.2 Facilitation of activities in reasoning-and-proving course ... 86
3.1.2 The Participants ... 86
3.2 DATA COLLECTION ... 89
3.2.1 Structured interviews ... 90
3.2.2 Course Notebooks ... 98
3.2.3 Artifact task packets ... 99
3.3 DATA CODING & ANALYSIS ... 101
3.3.1 Conceptions of Proof ... 102
3.3.1.1 An explanation of how the data was analyzed to address the four conceptions of proof? ... 105
3.3.2 Solving Reasoning-and-proving Tasks ... 107
3.3.3 Critiquing Student Arguments ... 111
3.3.4 Selecting Reasoning-&-Proving Tasks ... 116
4.0 CHAPTER 4. RESULTS ... 120
4.1 PRE-SERVICE TEACHERS CONCEPTION OF PROOF ... 121
4.1.1 Criteria of Proof: What is proof and what counts as proof? ... 123
4.1.1.1 Proof must be true for all cases ... 127
4.1.1.2 Claims and statements accepted by community ... 128
4.1.1.3 Type, form, and representation of a proof can vary ... 129
4.1.1.4 Summary of participant changes with respect to criteria of proof ... 131
4.1.2 Purpose: Why teach proof? ... 134
4.1.2.1 Purpose of proof before the start of the course ... 134
4.1.2.2 Purpose of proof during and after the conclusion of the course ... 137
4.1.3 Equity: Who should write proofs? ... 142
4.1.3.1 Participant perception of equity and proof writing: Prior to start of course ... 143
4.1.3.2 Participant perception of equity and proof writing: During and after the course ... 147
4.1.4 Opportunities: When (how often) should students engage in reasoning-and-proving activities? ... 152
4.1.4.1 What secondary courses should include reasoning-and-proving tasks? ... 153
4.1.4.2 How often will the participants include reasoning-and-proving activities? ... 156
4.1.5 Summary of participants' conceptions of reasoning-and-proving ... 161
4.2 PRE-SERVICE TEACHERS ABILITY TO CONSTRUCT PROOFS ... 163
4.2.1 Analyzing participant results ... 164
4.2.2 Reasoning-and-proving task analysis ... 169
4.2.2.1 N^2+N is always even ... 169
4.2.2.2 Squares Problem ... 173
4.2.2.3 Odd plus Odd is Even ... 176
4.2.2.4 N-by-N Window Problem ... 178
4.2.2.5 Parallelogram Construction Task ... 183
4.2.2.6 Sticky Gum Problem ... 187
4.2.2.7 Explaining Number Patterns Task ... 190
4.2.2.8 Calling Plans Task ... 192
4.2.3 Summary of participant challenges with reasoning-and-proving ... 195
4.2.3.1 Defining terms and variables ... 195
4.2.3.2 Transitioning from inductive to deductive reasoning ... 196
4.2.3.3 Stating conjectures ... 196
4.2.3.4 Clearly explaining thinking ... 197
4.2.3.5 Identifying a counterexample to prove a statement ... 197
4.3 PRE-SERVICE TEACHERS ABILITY TO ANALYZE STUDENTS REASONING-AND-PROVING SOLUTIONS ... 198
4.3.1 General analysis of student solutions and participant results ... 198
4.3.2 Analysis of the student solutions and participant results ... 200
4.3.3 Participant understanding of applying criteria of proof ... 203
4.3.3.1 Student solutions, which were straightforward to identify ... 204

4.3.3.2 Student solutions, which were more challenging to identify 206
4.3.4 Summary of participant growth with critiquing reasoning-and-proving solutions..... 229
4.4 PRE-SERVICE TEACHERS ABILITY TO SELECT AND OR MODIFY REASONING-AND-PROVING TASKS.... 231
4.4.1 Ability to modify tasks to include reasoning-and-proving opportunities 234
4.4.1.1 High-level make a generalization task 238
4.4.1.2 High-level provide an argument tasks 243
4.4.1.3 Low-level reasoning-and-proving tasks....245
4.4.1.4 High-level non-reasoning-and-proving 246
4.4.1.5 Summary of tasks selected and modified for the two course assignments....248
4.4.2 Five selected tasks for interview three 250
4.4.3 Ten implemented classroom reasoning-and-proving tasks.... 252
4.4.3.1 Karen's enacted proof lessons....252
4.4.3.2 Katie's enacted proof lessons....254
4.4.4 Summary of selecting and modifying reasoning-and-proving tasks 255

5.0 CHAPTER 5: DISCUSSION.... 257
5.1 IMPORTANCE OF STUDY: EXPLAINING THE RESULTS.... 257
5.1.1 Expanded conception of proof.... 258
5.1.2 Constructing proofs.... 261
5.1.3 Critiquing arguments.... 266
5.1.4 Selecting and or modifying reasoning-and-proving tasks 268
5.2 INTERESTING INSIGHTS THAT EMERGED ACROSS THE RESEARCH QUESTIONS 270
5.2.1 Comparing what they said with what they did.... 271
5.2.1.1 Conception of criteria versus critiquing students' solutions....272
5.2.1.2 Conception of purpose versus selecting and or modifying tasks....273
5.2.1.3 Conception of equity and opportunity versus implementing reasoning-and-proving tasks 274
5.2.2 Comparing ability to construct proofs versus critiquing arguments 276
5.3 CONCLUSION AND FUTURE RESEARCH IDEAS.... 277

APPENDIX A.... 281

APPENDIX B.... 296

APPENDIX C.... 314

BIBLIOGRAPHY 317

LIST OF TABLES

TABLE 3.1 COLLECTED DATA FOR ANALYSIS 89
TABLE 3.2. OPENED-ENDED INTERVIEW QUESTIONS 92
TABLE 3.3. REASONING-AND-PROVING TASKS: PART 2 OF EACH INTERVIEW 94
TABLE 3.4. ARGUMENT TYPES FOR EACH STUDENT SOLUTION: PART 3 OF EACH INTERVIEW 96
TABLE 3.5. CONCEPTIONS OF PROOF CATEGORIES 103
TABLE 3.6. REASONING-AND-PROVING TASK CODING TOOL 108
TABLE 3.7. TWELVE STUDENT SOLUTION DISTRACTER ITEMS 112
TABLE 3.8 TWELEVE DISTRACTER ITEMS REPRESTED ACROSS FIVE STUDENT WORK ACTIVITIES 114
TABLE 4.1. FOUR CONCEPTIONS OF PROOF CATEGORIES 122
TABLE 4.2. PARTICIPANTS' CRITERIA OF PROOF COMPARED AGAINST COURSE CRITERIA. 125
TABLE 4.3. PARTICIPANTS' BELIEFS ABOUT WHICH SECONDARY COURSES SHOULD INCLUDE PROOF 154
TABLE 4.4. PARTICIPANTS THINKING ABOUT HOW OFTEN THEY PLAN TO IMPLEMENT PROOF TASKS 157
TABLE 4.5. RESULTS OF PARTICIPANT SOLUTIONS TO ALL EIGHT R&P TASKS USING THE R&P CODES 166
TABLE 4.6. PERCENTAGE CORRECT WHILE EVALUATING REASONING-AND-PROVING SOLUTIONS 199
TABLE 4.7. TOTAL PARTICIPANT RESPONSES TO EACH OF THE 32 SOLUTIONS 202
TABLE 4.8. TWELVE PRE-DETERMINED DISTRACTERS 208
TABLE 4.9. TOTAL NUMBER OF REASONING-AND-PROVING TASKS EACH PARTICIPANT SELECTED 234
TABLE 4.10. RESULTS OF INITIALLY CHOSEN REASONING-AND-PROVING TASKS 235
TABLE 4.11. THE 18 SELECTED AND MODIFIED REASONING-AND-PROVING COURSE TASKS 236
TABLE 4.12. PARTICIPANTS' SKILL WITH SELECTING AND MODIFYING REASONING-AND-PROVING TASKS 249
TABLE 5.1. PURPOSES OF PROOF THAT THE PARTICIPANTS IDENTIFIED 273

LIST OF FIGURES

FIGURE 2.1. ERIC'S ANSWER........ 27
FIGURE 2.2. CONJECTURE MADE IN MR. D'S CLASS........ 35
FIGURE 2.3. CONTENTS OF DAHLBERG AND HOUSMAN INTERVIEW........ 39
FIGURE 2.4. REASONING-AND-PROVING FRAMEWORK ADAPTED FROM G. STYLIANIDES (2010)........ 47
FIGURE 2.5 COGNITIVE LEVELS OF PROOF TASKS........ 49
FIGURE 2.6. MATHEMATICAL TASKS FRAMEWORK ADAPTED FROM (STEIN & LANE, 1996) 52
FIGURE 2.7. UNPACKING PHASE 1 OF MTF AND IDENTIFYING FACTORS THAT INFLUENCE TASK MODIFICATION........ 79
FIGURE 3.1. COURSE MAP SUMMER 2011........ 85
FIGURE 3.2. DATA COLLECTION TIMELINE........ 90
FIGURE 3.3. TANYA'S SOLUTION TO THE $N^2 + N$ IS ALWAYS EVEN TASK........ 110
FIGURE 3.4. STUDENT B SOLUTION TO THE "O + O = E" TASK........ 114
FIGURE 3.5. SPECTRUM OF SOLUTIONS TO REASONING-AND-PROVING TASKS........ 115
FIGURE 4.1. N^2+N IS ALWAYS EVEN TASK........ 169
FIGURE 4.2. THE SQUARES PROBLEM........ 174
FIGURE 4.3. ODD + ODD IS EVEN TASK........ 176
FIGURE 4.4. N-BY-N WINDOW PROBLEM........ 179
FIGURE 4.5. PARALLELOGRAM CONSTRUCTION TASK........ 184
FIGURE 4.6. THE STICKY GUM PROBLEM........ 188
FIGURE 4.7. EXPLAINING NUMBER PATTERNS TASK........ 191
FIGURE 4.8. THE CALLING PLANS TASK........ 193
FIGURE 4.9. STUDENT B'S SOLUTION TO $N^2 + N$ IS ALWAYS EVEN........ 204
FIGURE 4.10. STICKY GUM STUDENT SOLUTION F........ 205
FIGURE 4.11. TYPES OF ARGUMENTS THAT WERE DISTRACTERS........ 207
FIGURE 4.12. STUDENT'S D SOLUTION TO $N^2 + N$ IS ALWAYS EVEN........ 209
FIGURE 4.13. STUDENT I SOLUTION TO THE "O + O = E" TASK........ 210
FIGURE 4.14. STUDENT E'S SOLUTION TO $N^2 + N$ IS ALWAYS EVEN........ 213
FIGURE 4.15. STUDENT B SOLUTION FROM THE "O + O = E" TASK........ 215
FIGURE 4.16. STUDENT A SOLUTION TO NXN WINDOW........ 219
FIGURE 4.17. STUDENT C SOLUTION TO THE STICKY GUM PROBLEM........ 221

FIGURE 4.18. STUDENT A SOLUTION TO THE CALLING PLANS TASK 223
FIGURE 4.19. STUDENT SOLUTION E TO THE CALLING PLANS TASK 226
FIGURE 4.20. TIMELINE OF TASK SELECTION AND MODIFICATION 232
FIGURE 4.21. COURSE DEVELOPED PRINCIPLES FOR REASONING-AND-PROVING TASK MODIFICATION 237
FIGURE 4.22. KAREN'S INITIAL TASK AND MODIFIED VERSION COMPLETED HER SECOND COURSE ASSIGNMENT 240
FIGURE 4.23. TINA'S EXPONENT TASK THAT IS HIGH-LEVEL AND MAKE A GENERALIZATION 241
FIGURE 4.24. LUCY'S SELECTED AND MODIFIED PARALLEL LINES PROBLEM 242
FIGURE 4.25. KATIE'S ASSIGNMENT ONE PROOF TASKS 244
FIGURE 4.26. NATHANIEL'S MODIFIED AREA OF TRIANGLE TASK 245
FIGURE 4.27. BRITTANY'S PARABOLA PROBLEM, WHICH STAYED A LOW-LEVEL PROVIDE AN ARGUMENT TASK 246
FIGURE 4.28. TANYA'S PARABOLA PROBLEM 247
FIGURE 4.29. KATHERINE'S QUADRATIC FUNCTION PROBLEM 248
FIGURE 4.30. KAREN'S BLOCKS TASK 251
FIGURE 4.31. KATIE'S SQUARES AND SHADED AREA PROBLEM 252

ACKNOWLEDGEMENTS

Finishing this dissertation is a grand accomplishment for me. However, I believe no great feat is attained alone. I received an abundant amount of support and insightful feedback along this journey from many individuals, in particular, my committee members: Dr. Peg Smith, Dr. Ellen Ansell, Dr. James Greeno, and Dr. Charles Munter. I want to first and foremost thank my advisor, Dr. Peg Smith. I really appreciated the multiple rounds of feedback you provided on each chapter to get it to its current form. I look forward to continue collaborating with you on our current project and future adventures. You have been more than an advisor to me; I will always remember you as the greatest teacher I ever had. I feel very blessed to have had the opportunity to work with you over the past six years.

Ellen, I thank you for your detailed insights on the organization of the chapters and helpful suggestions to improve the clarity of the content. Chuck, I really appreciate you pressing my on the design of the study from a research perspective and future considerations about making connections to learning trajectories. And I always want to give a big thanks to Jim. I thank you for taking the time to help me better understand the differences and commonalities among several learning theories, in addition to the direction you provided towards thinking about research design. Without all of the thoughtful feedback from all of my committee members, this final product would not have been the same. Most importantly, I am grateful for the insights regarding future directions for this research and where more work is needed.

I would also like to thank my fellow graduate students: Sam, Adam, Michelle, Scott, Mary Ellen, Josh, and Steve. I really appreciate the feedback you each provided on my coding tools and presentations related this work. Sam, I want to specifically thank you for your insights and acting as my reliability rater.

I also want to acknowledge my new University of New Mexico colleague, Vanessa, for her formatting support and gracious hospitality. This document would not have been uploaded on time without you, so thank you.

Last, but definitely not least, I need to thank my favorite person and partner, Anne-Marie. Thank you for reading various portions of this document and dealing with me as I isolated myself to work so many nights and mornings on my dissertation. I thank you for understanding and supporting me through this journey.

1.0 CHAPTER 1: STATEMENT OF THE PROBLEM

1.1 INTRODUCTION

Proof is the foundation of mathematics. Unlike the sciences, where truth is based on tested trials to make claims about a larger population, mathematical proof provides truth for all cases beyond specific cases. Mathematicians use deductive reasoning to explain why a situation is always true throughout all areas of the discipline. School mathematics, in contrast, has historically relegated proof to a single high school geometry course while promoting a single axiomatic form, portraying a constrained view of this essential activity unique to mathematics. However, over the past decade, mathematics educators have recommended that proof become a more central activity across all elementary and secondary courses for all student ability levels (Ball, Hoyles, Jahnke, & Movshovitz-Hadar, 2002; Knuth, 2002a, 2002b; Sowder & Harel, 1998). In addition to increasing the access to proof in mathematics classrooms, there exists supporting research on productive instructional methods to learn proof construction (e.g. Lanin, 2005; Martin, McCrone, Bower, & Dindyal, 2005; A. J. Stylianides & G. J. Stylianides, 2009) beyond the conventional two-column form. While the suggestions for changing the handling of proof in schools has gained momentum as an integrated purposeful activity, the practical adjustment is challenging.

The curricula that schools adopt include limited opportunities for students to learn what proof is and how to construct valid arguments. Johnson, Thompson, and Senk (2010) discovered that less than six percent of tasks in high school textbooks outside of geometry prompt students to reason or prove mathematical situations. However, this curricular issue may now being addressed. Textbook publishing companies align the mathematics problems in the curricular materials with state adopted standards. A common set of mathematics standards were adopted by 45 states, include the development of argumentation across all grade levels and mathematics courses (CCSSM, 2010). However, simply adding proof activities in textbooks to align with the recently published standards is only a part of the practical problem.

The larger obstacle is paradoxical. If reasoning and proof activities have been all but absent from high school curricula, excluding geometry (Johnson, Thompson, and Senk, 2010), then it can be concluded that teachers have not been provided sufficient resources to enact such tasks. However, simply adding reasoning and proof tasks to the curricula is insufficient. Bieda (2010) observed experienced teachers implementing proof tasks and noticed that the teachers did not hold students responsible for justifying their thinking. Additionally, Knuth (2002a, 2002b) ascertained that many high school teachers misunderstood the meaning of mathematical proof and struggled to identify valid from invalid solutions. There in lies the conflict and need for teacher learning. If curricula materials align with the new standards, then the research community will continue to report that teachers lack knowledge of proof to enact such tasks effectively. Alternatively, fostering teacher knowledge of proof without sufficient resources in the curriculum could result in a continued near absence of proof instruction. Therefore, a practical solution is to expand teachers' knowledge of proof, along with skills to support their students' learning, while addressing the need to identify and or modify tasks within their resources to provide such opportunities. Learning how to foster student access and development of proof, along with focused exercises on task modification, may support teachers in providing rich proving opportunities for their students.

This design experiment[1] will investigate the impact of a curriculum project, Cases Of Reasoning and Proving in Secondary Mathematics (CORP) on pre-service teachers' learning of reasoning-and-proving in a Masters level mathematics methods course. The study described herein is the second implementation of the CORP materials. The findings from the first enactment suggest that the practicing teachers improved their ability to write proofs and identify valid student arguments. However, follow-up work revealed that the teachers struggled to select or modify worthwhile tasks so that they could provide opportunities for their students. Since the tasks in which students engage shape their thinking about the subject (Doyle, 1988; Stein, Grover, & Henningsen, 1996) and the teachers in the initial enactment of the CORP materials exhibited a limited ability to choose or modify reasoning-and-proving[2] tasks, the materials were redesigned to address this issue. Therefore, an emphasis of this second iteration was to deepen secondary teachers' mathematical knowledge for teaching reasoning-and-proving tasks in their classroom. The identified mathematical knowledge for teaching proof addressed in the course curriculum focused on advancing three areas: writing valid arguments, critiquing and questioning student thinking, and selecting and implementing appropriate student tasks. Additionally, this study investigates the participants' changes in conceptions of reasoning-and-proving for teaching.

The hypothesis driving this study is that through engaging in this second iteration of the reasoning-and-proving course for mathematics teachers, the participants will increase their

---

[1] Design experiments are meant to influence theory and practice through an iterative process of changing the design and researching the effects (Brown, 1992; Greeno, 2006). The research herein is the second iteration of the design.
[2] The hyphenated term reasoning-and-proving will be explained in more detail later in this chapter.

ability to identify valid arguments, improve their skill at constructing proofs, and become better prepared to select or modify tasks from their curricula. Moreover, while explicitly engaging in activities to increase their knowledge in each of the three specified areas, the participants will expand their conceptions of proof and the role it should play in secondary classrooms. The remainder of this chapter will provide further justification for why teacher learning is required to increase secondary students' opportunities to write and critique proofs.

1.2 BACKGROUND

The type of mathematical tasks in which teachers engage students, influence the type of learning students experience (Doyle, 1988; Stein, Grover, & Henningsen, 1996). Implementing tasks that provide opportunities for students to make sense of mathematical concepts requires teachers to have knowledge of content and how students will progress mathematically including the misunderstandings they might have (Carpenter et al., 1989; Ball Thames, & Phelps, 2008). Writing proofs requires students to justify why a conjecture is true and to convince their classmates, as well as the teacher, that it is true (Hersh, 1993). With the political backing of the National Council of Teachers of Mathematics (NCTM) (NCTM, 2000, 2010) and newly released Common Core State Standards in Mathematics (CCSSM) (CCSS, 2010), reasoning and argumentation are expected to be included in all K-12 courses. In order to make this vision a reality, teachers must gain an understanding of reasoning-and-proving for teaching so that they are capable of fostering student learning of these practices.

1.2.1 Policymakers Support Reasoning and Sense Making in Schools

The 1983 federal report: *A Nation at Risk* highlighted the dire need for educational improvement and specifically addressed mathematics as a major area of concern. Early drafts of the federal report provoked the National Council of Teachers of Mathematics (NCTM) to publish *An Agenda for Action* (NCTM, 1980). The NCTM message was that basic skills are taught at the expense of understanding. Instead of explicitly addressing proof in *An Agenda for Action*, the document promoted problem solving. In 1989, NCTM retreated from singly promoting problem solving and introduced *Curriculum and Evaluation Standards for School Mathematics* (CESSM). However, proof was under represented in CESSM. Only two of 14 CESSM standards listed proof and reserved it for high attaining students. While the 1989 content standards promoted mathematical understanding, it fell short of suggesting formal reasoning for all students across all secondary courses. As a result, it is logical to conclude that most schoolteachers and textbook publishers also did not pose proof tasks during this period.

In 2000, NCTM released *Principles and Standards for School Mathematics* (PSSM) and made a challenging statement:

> *Reasoning and proof are not special activities reserved for special times or special topics in the curriculum but should be a natural, ongoing part of classroom discussions, no matter what topic is being studied (NCTM, 2000, p. 342).*

This was and still may be provocative since it directly confronts the conventional treatment of proof in school mathematics. The document went beyond making the proclamation to include reasoning and proof as one of five process standards and includes four expectations for all students:

- Recognize reasoning and proof as fundamental aspects of mathematics
- Make and investigate mathematical conjectures
- Develop and evaluate mathematical arguments and proofs

- Select and use various types of reasoning and methods of proof

(NCTM, 2000)

The PSSM document was the first to recognize the critical role proof plays in school mathematics.

NCTM most recently published a series of reasoning and sense making books. The first book was published in 2009 titled: *Focus in High School Mathematics: Reasoning and Sense Making.* Three follow-up books concentrated on specific content areas: statistics and probability, algebra, and geometry. Reasoning encompasses a variety of activities such as: explaining, investigating, making conjectures, and deductive argumentation that are all construed from assumptions and or definitions. Sense making involves examining contexts and linking it with prior knowledge. The authors define reasoning and sense making as a twisting thread in which the two are interconnected and move along a spectrum from informal to formal mathematical justification where formal reasoning and sense making both include proof (NCTM, 2009).

In June 2010, a monumental political shift occurred in education. Historically, local districts or states controlled curriculum standards. In the summer of 2010, selected mathematicians and mathematics educators worked with the National Governors Association to publish a document titled *The Common Core State Standards* (CCSS). As of August 2012, 45 states plus the District of Columbia and the US Virgin Islands had formally adopted the CCSS. These standards will be reflected in the school curricula and assessments for the adopted states, resulting in the potential to directly impact instruction. Teachers will be provided textbooks that are aligned with the CCSS and over the next few years, student assessments are expected to mirror the new standards as well.

According to *The Common Core State Standards for Mathematics* (CCSSM), "One hallmark of mathematical understanding is the ability to justify, in a way appropriate to the student's mathematical maturity, why a particular mathematical statement is true or where a mathematical rule comes from" (CCSSM, 2010, p.3). Similar to the 2000 PSSM recommendations, the CCSSM identifies mathematical processes. While proof is not explicitly listed in CCSSM, the full spectrum of informal and formal reasoning is stated. For instance, the second mathematics practice standard reads as the following: *Reason abstractly and quantitatively.* A main thrust here is to build fluency with connecting problem context with generalizations. The third math process also directly relates to proof, *Construct viable arguments and critique the reasoning of others*. The seventh practice standard lies on the reasoning and sense-making spectrum as well: *Look for and make use of structure.* Here students are expected to understand and make connections across equivalent mathematical expressions or objects. The descriptions of these three mathematical practices along with the five others are in line with the view of proof as a communal activity while promoting a focus on conceptual understanding (Bell, 1976; Hanna, 1995).

Now that national mathematics education policy is in agreement with mathematics educators' view on the importance of proof in the classroom, other variables must be addressed. While establishing standards is encouraging, standards alone are not enough to integrate proving opportunities across all secondary courses and classrooms. Two questions need to be addressed: 1) What is proof and what are the recommendations for it in secondary mathematics? 2) What is needed to support students' learning of proof?

1.2.2 What is proof? And what are the recommendations for it in secondary mathematics?

Defining mathematical proof is different from identifying characteristics for judging the validity of a presented argument. For instance, Hersh defines a proof as a "convincing argument, as judged by qualified judges" (1993, p. 389). This obtuse definition is consistent across reports on proof in mathematics education (Reid, 2005). Determining what convinces qualified judges is subjective based on the community in which the argument is presented (Harel & Sowder, 2007; Polya, 1945; Reid, 2005; A.J. Stylianides, 2007). Instead of focusing on an agreed upon definition, researchers have identified characteristics of arguments that qualify as proof (A.J. Stylianides, 2007; Weber, 2008). The point is that a definition of proof does not necessarily translate into accurately judging solutions to the extent to which they prove.

A.J. Stylianides (2007) developed a criterion for judging arguments based on the context in which the argument is produced as listed below:

> *Proof is a mathematical argument, a connected sequence of assertions for or against a mathematical claim, with the following characteristics:*
>
> 1. *It uses statements accepted by the classroom community (set of accepted statements) that are true and available without further justification;*
> 2. *It employs forms of reasoning (modes of argumentation) that are valid and known to, or within the conceptual reach of, the classroom community; and*
> 3. *It is communicated with forms of expression (modes of argument representation) that are appropriate and known to, or within the conceptual reach of, the classroom community. (2007, p. 291)*

He argues that each classroom community should develop a list of appropriate characteristics to foster student understanding of proof. Teachers could use these dimensions to co-construct a criterion of what counts as proof in their classroom.

Hanna (1990) and others (e.g. Bell, 1976; Hersh, 1993) argue that proof in schools should *explain* the truth or fallacy of a conjecture. Hanna contends that formal logical arguments such as ones following the mathematical induction method are proofs, but such prescribed procedures fail to help students understand the validity of the claim. She writes, "Proofs by mathematical induction are non-explanatory in general" (Hanna, 1990, p. 10). Explanatory power can be found through connecting algebraic symbols with diagrams where the diagram can help students make sense of why the claim is true or false. Harel and Sowder (2007) classify a non-explanatory proof as one that follows the external conviction schemes where a student only follows the methods described in textbooks or those completed by a teacher without understanding the mathematical concepts questioned in the assertion to be proven. The argument is that students are capable of deductive reasoning; however, the form and organization of the reasoning should be aligned with students' current thinking so as to not force a formal structure.

1.2.3 Students' Ability to Prove

Proof, when taught in schools, has consistently been an activity in which students at all levels of education struggle, both in the United States and internationally (Bell, 1976; Healy & Hoyles, 2000; Recio & Godino, 2001; Porteous, 1990; Senk, 1985). First and foremost, students are not qualified judges. They are easily convinced that an assertion is true, even if the argument falls short of justifying all cases (Boaler & Humphries, 2006; Chazan, 1993; Healy & Holes, 2000; Hersh, 1993; Porteous, 1990). While it is readily apparent that students struggle to write and critique proofs, the following studies aid in unpacking what particular issues students encounter with the activity.

Bell (1976) placed solutions written by 15 year-old students from the United Kingdom into two groups, empirical and deductive arguments, with four parallel levels described for each argument type. The levels ranged from a failure to work with the conditions of the problem to constructing a proof. For instance, the spectrum of solutions in the empirical group include a student generated example that does not relate to the situation (lowest empirical level) and exhausts all examples in the finite problem set for a complete proof (highest empirical level). Bell reported on two problems with 32 responses to each question. One problem, which involved divisibility of three, allowed for students to exhaust all 14 cases and this was the only proof method chosen. Almost one-fifth of the students chose this method meaning that 80% did not construct a proof. Bell was surprised that none of them checked a few cases and then generalized a pattern to account for all possible numbers or used algebra at all. About half of the students' solutions showed a misinterpretation of the question or concept. The second problem required a proof by constructing triangles to meet a set of given conditions. Almost 40% of the students misunderstood the term congruent or the problem altogether. No student wrote a correct proof for the second problem. So in both cases, the majority of students were unable to write proofs since they did not understand the question or the mathematical content, even though Bell explicitly chose tasks he thought were content accessible for the age group.

Senk (1985) analyzed US students' ability to write proofs. She reviewed 1520 students' solutions from eleven schools in five states, which was a subset of the data collected by the Cognitive Development and Achievement in Secondary School Geometry (CDASSG) Project. All of the students in the study were enrolled in a geometry course in which proof was part of the instruction. Almost one-quarter (24%) of the students were in honors classes, a little less than half (46%) were in classes labeled regular, and the final 30% were in heterogeneously grouped classes. The classroom teachers administered the test during the last month of the school year. The six test items resembled those commonly found in geometry textbooks and varied in difficulty level. The students' responses were scored using a rubric ranging from zero (writes little or nothing) to four (writes a proof with at most one error in notation). Overall, 3% scored a perfect score on each of the six items and 29% did not write a single valid proof. Senk concluded that about 25% demonstrated zero competency, 25% could write trivial proofs, 20% could do some proofs with complexity, and 30% mastered constructing proofs as presented in common geometry curriculum. So even after enrolling in a geometry proof writing course, the overwhelming majority of the students did not learn the skill.

Recio and Godino (2001) asked two large groups of students entering the University of Córdoba (Spain) to write two proofs during the 1994-95 and 1997-98 school years. In 1994, 429 students worked on the two problems and in 1997 the same two questions were given to 193 students. All of the students were enrolled in a university mathematics course. The following task, based on number theory, was given to all students: *Prove that the difference between the squares of every two consecutive natural numbers is always an odd number, and that it is equal to the sum of these numbers.* The second task involved elementary geometry content: *Prove that the bisectors of any two adjacent angles form a right angle.* The researchers provided the students with definitions for both questions including natural numbers, bisector, adjacent angle, and right angle in order to support their ability to write proofs based on a possible lack of content knowledge. The responses for both questions were arranged along five levels (a rubric) starting at incoherent and ending at proof. The middle levels included empirical examples and generalizations with partially correct procedures. Recio and Godino found in 1994-95 that less than 50% of the students wrote a correct proof for either problem and only 32.9% of the students wrote a proof for both problems. In 1997-98 only 22.8% of the students wrote a proof for both mathematics statements. So even though the mathematical concepts were elementary for the university mathematics students, the majority of them were unable to produce deductive arguments.

While this sample does not exhaust the research, it highlights an ongoing problem: The majority of high school students lack the skills needed to write convincing valid arguments. Therefore, students need a different level of support and more opportunities to develop their skills in writing and critiquing proofs in order to meet the recommendations of mathematics educators and policymakers.

1.2.4 Treatment of Proof in School

From both a pedagogical and learning perspective, the transition to deductive reasoning is abrupt (Moore, 1994). The problem with the conventional handling of proof in school mathematics is that when addressed, it is presented as a completed product (Chazan, 1990). Starting with formal deductive logic hinders students' intuitive reasoning skills and distorts the purpose of proof (Ball et al., 2002; Hanna, 1990). Many students are frustrated when presented with a proof task, since they do not understand where to start (Moore, 1994; Solomon, 2006). Teachers often present a single authoritative proof argument based on a given conjecture for students to simply memorize and reproduce (Harel & Sowder, 1998; Harel & Rabin, 2010).

Proof is usually organized in mathematics classrooms as a ritual without meaning (Ball et al., 2002). The typical focus is on systemization and strategy, but instead of allowing students to choose how to organize their argument, they are generally expected to follow a particular form (i.e. two-column proof). Students are asked to verify known facts, as opposed to searching for their own conjectures to verify. Starting with formal methods and mundane exercises paints a contorted picture of the purposes of proof. Furthermore, supplying the two-column formal deductive strategy eliminates the opportunity to develop new techniques for solving novel problems. Rigor needs to play a secondary role to understanding (Ball et al., 2002; Hanna & Jahnke, 1996). However, Coe and Ruthven (1994) claim the opposite (rigor over understanding) has been the case, since the writing on proof has been from a philosophical rather than pedagogical stance. Implementing proof from a philosophical stance has shown to be unsuccessful, starting with the "new math" era in the 1960s (Hanna, 1995). Hanna argues the major challenge to integrating proof throughout the curriculum is finding more ways to use it to promote mathematical understanding. Therefore, proof needs to become a more substantial part of mathematics education, since it has the potential to deepen students' understanding. However, teachers need to learn to identify opportunities within the curriculum to expose students to proof in authentic ways.

G. Stylianides's (2008) Reasoning-and-Proving framework provides access for learners (students and teachers) to engage in the work of developing proofs modeled after the practice of mathematicians. The hyphenated combination of reasoning-and-proving is intentional, representing a specific connotation, and will be utilized throughout this study. Reasoning consists of searching for patterns and proposing a conjecture from observations. Proving involves constructing proof or non-proof arguments to justify the generalization. Therefore, the combination of reasoning-and-proving implies a set of activities, which starts with searching for a pattern or making a conjecture, and finishes with constructing an argument to justify why the proposed conjecture is always true. While this could be interpreted as a clean or strict linear process, it is quite the opposite. Students may start by generating examples or make observations and settle on a conjecture, only to find a counterexample. The discovery may cause them to look for a different conjecture for the situation. This process of allowing students to struggle with a mathematical situation before formalizing a proof has led to promising results (Healy & Hoyles, 2000; Lannin, 2005; Smith, 2006).

1.2.5 Teachers' Ability to Write Proof

A plausible reaction to students' limited ability to write and critique proofs might be to have teachers spend more time teaching proof. This thinking aligns with the recommendations prescribed in the Common Core State Standards for Mathematics (2010) and NCTM (2000) standards and also supports the thinking of mathematics educators (Ball et al., 2002). A problem with this suggestion is that teachers, like students, are unable to critique or write proofs on a consistent basis.

Knuth (2002a) collected semi-structured interview data on 16 practicing high school mathematics teachers. For part of the interview, he presented teachers with researcher-produced solutions that varied with respect to validity. He focused on presenting arguments that were mathematically accessible, so the content did not serve as an obstacle to understanding the argument. They were given five problems, each with three to five solutions, for a total of 21 responses (13 proofs and 8 non-proofs) to evaluate. Each set of five statements contained at least one argument that was not a proof. Teachers were asked to use a scale of one (non-proof) to four (proof) to rate the 21 arguments. A score of two or three provided teachers an opportunity to discuss issues such as assumed truths and completeness. The results indicated that the 16 teachers were successful at identifying the 13 proof arguments (93% correct). On the other hand, only two-thirds of the non-proof arguments were labeled as such. Additionally, every teacher labeled at least one non-proof as valid. Knuth concluded that the teachers judged the statements as correct or agreed with the truth of the statements based on certain non-essential criteria. In other words, he claimed that teachers lack a complete view of the criteria necessary for validating a proof.

In another study, Selden and Selden (2003) examined eight, undergraduate mathematics (4) and mathematics education (4) majors' ability to validate arguments. During an interview, students were presented with four solutions to the same mathematics statement: *For any positive integer n, if* n^2 *is a multiple of 3, then n is a multiple of 3*. The interviews lasted about an hour and contained four parts. The first asked students to think about and make sense of the statement and to write a proof if possible. Only two of the eight students wrote a proof. The second phase of the interview asked them to read through each of the four solutions individually and think aloud while reading and reasoning through the argument. The third section presented the interviewees with all four arguments on the same page and asked them to make a decision as to the validity of the solution. If not a proof, they were to explain what statements in the argument were incorrect. The final section of the interview asked the students to explain how he or she judged the validity of an argument. The researchers recorded four different times (time 1 through time 4) during phases 2 and 3 of the interview in which students judged each argument. The four time periods allowed students to change their mind as they engaged in dialogue about specific statements in the solution with the interviewer. During time one, the eight students judged less than half (46%) of the four solutions correctly. However, the percentage increased to 81 at time four when students were asked to make their final decision about the solution. The researchers conclude that at the initial viewing, mathematics and mathematics education students judged arguments no better than chance, but with explicit attention to structure and less attention on surface features, the students learned to identify proofs.

College graduates with degrees in mathematics enter credentialing programs to become high school mathematics teachers. So if university undergraduate mathematics courses are the only opportunity where teachers learn to write proofs, and research shows that some students in such courses struggle, then it is fair to conclude high school math teachers need a different experience constructing and learning about proofs.

Appropriating more instructional time for skills with which teachers themselves struggle is not likely to improve their students' understanding. In addition to Hanna's (1995) challenge that more opportunities need to be added into the existing curriculum for students to reason and prove, teachers also need to learn to write and critique arguments. Even if student proof activities existed, many teachers are not qualified to support student learning, which is why students continue to demonstrate poor results when their ability to write proofs is assessed. Therefore, there is a need in mathematics education to advance teachers' ability to construct and critique proof arguments.

1.2.6 Teachers and Students Harbor Deep Misconceptions of Proof

Since proof has often been presented as a meaningless isolated activity, it should not be surprising that learners misunderstand the role and purpose it plays in mathematics. Students and teachers are convinced mathematical statements are true without proof and do not think of proofs as convincing arguments (Coe & Ruthven, 1994; Chazan, 1993; Housman & Porter, 2003; Knuth 2002a; Martin & Harel, 1989). In other words, proof has become simply an institutional exercise, students and teachers engage in without learning its central purpose in mathematics (Solomon, 2006).

Chazan (1993) interviewed high school students at multiple schools and discovered a variety of misunderstandings. Some thought examples were proof. Some students in his study believed proofs themselves were not completely convincing, or that a counterexample could exist even for an argument they considered to be a proof. For instance, one of the arguments Chazan presented students contained a set of four individual cases. Some labeled the four-example argument as non-proof, but were still convinced that the statement was true. Another interview question included a typical geometry textbook two-column proof. Most interviewees identified it as proof, but claimed the argument only held for the provided diagram. In other words, the proof did not represent a general case. While it is important to engage students in more reasoning-and-proving activities to improve mathematical understanding, they also need to understand the implications of mathematical proof.

High school math teachers also misunderstand the meaning of proof. Knuth (2002b) learned that even though teachers could identify proof arguments, they were more convinced by empirical examples and non-proof arguments. Bieda (2010) found secondary teachers at the middle school level did not press students to produce a convincing general argument that held true for all cases, even when the activity asked for proof. If teachers understand a set of generated examples is not a proof, but find them convincing for a particular mathematical situation, then they are inclined to accept an incomplete justification, which sends students an incorrect message about what is needed to validate a conjecture. Secondary mathematics teachers need to learn a criterion of proof to establish a classroom community understanding of what constitutes a proof, which would then be used to critique presented student solutions (Knuth, 2002b; A. Stylianides, 2007). Therefore, in addition to expanding teachers' ability to write valid arguments, their conceptions of the meaning of proof should be examined in order to resolve possible misunderstandings so that they are capable of generating a classroom criterion that can be used to critique student written solutions.

1.2.7 Conclusion

Since proof has been all but absent from high school courses outside geometry, current curriculum is lacking in reasoning-and-proving activities (Hanna, 1995; Johnson, Thomson, & Senk, 2010). Hanna (1995) explains that proof activities should be integrated in the curriculum across high school courses, blending with and enhancing the curriculum teachers are using. Geometry proof tasks would need to be altered too. Most reasoning-and-proving activities in geometry only prompt for a proof rather than providing opportunities for students to search for patterns or make their own conjectures. All students in all high school courses should be provided opportunities to solve tasks that cover the full spectrum of reasoning-and-proving activities to learn mathematics and understand that proof is not a static object to be memorized (Ball et al. 2002; CCSS, 2010; Housman & Porter, 2003; Harel & Sowder, 1998; Smith, 2006; NCTM, 2000). This implies teachers will need to expand their knowledge of reasoning-and-proving beyond solving and critiquing arguments, but also to identify opportunities within their current curricula to access and advance their students' thinking.

1.3 POLICY, PROFESSIONAL DEVELOPMENT, AND THE TEACHER COURSE

Policymakers promote greater student expectations, and believe the catalyst for change is more ambitious instruction (Darling-Hammond, Wei, Andree, Richardson, & Orphanos, 2009; Spillane & Jennings, 1997; Stigler & Hiebert, 2004; USDE, 2000). The shift in ideology to increasing student learning outcomes is to build teacher autonomy and community (Darling-Hammond et al., 2009). Prescribing curriculum guides and standards from districts or state departments without sustained content focused teacher learning has shown to be unproductive (Cohen, 1990; Firestone, Mangin, Martinez, & Polovsky, 2005; Spillane & Zeuli, 1999). Teachers require professional development opportunities and teacher education courses to expand their knowledge of content, students' thinking, and pedagogical methods so they are prepared to make autonomous decisions, which are connected to teaching practices that promote student understanding (Ball & Cohen, 1999; Borko, 2004; Thompson & Zeuli, 1999).

Hiebert and Grouws (2007) define teaching as: "… classroom interactions among teachers and students around content directed toward facilitating students' achievement of learning goals" (p. 372). Ball and Cohen (1999) explain that professional teacher learning needs to be directly connected to the work of teaching. This implies that teachers should be provided opportunities outside the classroom to practice the work they are being asked to conduct in their classrooms. Therefore, professional developers need to engage teachers in learning situations that closely mirror the intended work of classroom teaching.

The NSF funded teacher curricular materials titled: *Cases Of Reasoning and Proving in Secondary Mathematics* (CORP), which form the basis of the course that is the focus of this study, exemplifies these suggestions for best practice. The activities focus on learning about reasoning-and-proving (the content), how students construct arguments, and ways to instruct and integrate proof into practice. For example, in the second unit teachers are asked to write a proof analyze student solutions to the same task, and finally to consider possible questions to pose to students whose solutions fall short of proof. In other words this sequence first provides teachers an opportunity to write a proof, consider how their students might engage in the same task, and finish by thinking about questions to support students' current thinking and other questions to move student thinking toward a learning goal. A course objective is to build teacher knowledge of reasoning-and-proving so that teachers are well prepared to make autonomous curriculum decisions and the teachers can prepare students to develop a complete view of proof, which is connected to the larger mathematics community.

1.4 PURPOSE OF THIS STUDY

The purpose of this research study is to investigate the impact of a Masters level mathematics teacher education course that focuses on expanding teachers' knowledge of reasoning-and-proving on pre-service secondary teachers ability to: (1) write valid arguments, (2) critique and question presented solutions, and (3) select and implement reasoning-and-proving tasks for their students.

1.5 RESEARCH QUESTIONS

This study analyzes pre-service teachers learning from a six-week course on reasoning-and-proving. The course involved the participants in solving problems, analyzing solutions, considering narrative case studies, and connecting the course activities with their curriculum. Additionally, sets of tasks, student solutions to the tasks, and teacher artifacts were collected as the participants' transition into their first year of teaching secondary mathematics. In particular, the study examines the following questions:

1. How do pre-service teachers' conceptions (i.e. purpose of proof, what counts, proof in secondary courses) of proof change over the duration of a course focused on reasoning-and-proving?

2. To what extent do pre-service teachers construct valid and convincing arguments when prompted to write proofs over the duration of a course focused on reasoning-and-proving?

3. To what extent do pre-service teachers improve their ability to distinguish between proof and non-proof arguments created by students over the duration of a course focused on reasoning-and-proving?

4. To what extent do pre-service teachers improve their ability to select and or modify reasoning-and-proving tasks for students over the duration of a course focused on reasoning-and-proving and during their first year in the classroom?

1.6 SIGNIFICANCE

This study hypothesizes that improving secondary students' ability to reason-and-prove is based on the knowledge and opportunities provided by their classroom teachers. If teachers can prove mathematical situations, critique and question their students' thinking, and select and implement reasoning-and-proving tasks, then their students will improve their skill at writing and analyzing arguments. The main purpose of the study is to determine the extent to which the teachers' knowledge changes throughout the course and the impact this has on their practice during their first year as practicing teachers. The results of the study contribute to the research knowledge base on teacher education. Thus it has the potential to identify activities that improve teacher knowledge of reasoning-and-proving in addition to the potential to identify at a smaller grain, challenges teachers face as they engage in this work.

This research is also expands on the Stylianides, G. J. and Stylianides, A. J. (2009) coding system to validate arguments. The coding tool developed for this study includes sub-codes for each of the main argument types while also incorporating the idea of clear and convincing statements. The statements are applied to valid arguments to distinguish among those that clearly state terms, define variable, and include a conclusion from arguments that are less convincing. The coding system described in this study could have the potential to validate solutions constructed in any K-16 mathematics course including teacher education.

1.7 LIMITATIONS

This study has several limitations. First, the sample is based on convenience choosing subjects who were admitted to a certificate program at a tier-one research university. Additionally, the sample is small, consisting of only nine participants. Thirdly, these nine teachers previously engaged in math methods courses, which focused on similar activities to enhance their knowledge of teaching. For instance, the subjects had prior experiences studying student solutions, analyzing narrative cases, and solving math tasks in several ways. Finally, the instructor of the course was a novice teacher educator and is a member of the curriculum development team. So a more experienced facilitator or one that is less familiar with the course materials may generate different learning outcomes. Hence, these results may not generalize to teachers more broadly.

1.8 OVERVIEW

This document is organized into five chapters. Chapter One argues the need for better preparing teachers who can to enact reasoning-and-proving tasks in their classrooms. Chapter Two reviews previously conducted research of proof, while focusing on related frameworks and theories on teacher learning and knowledge needed to teach reasoning-and-proving. Chapter Three describes the methodology including the data sources and analysis procedures used in this study. Chapter Four identifies the results of the analysis. Chapter Five presents the discussion of the findings, conclusions, and outlines suggestions for future research.

2.0 CHAPTER 2: REVIEW OF LITERATURE

The purpose of this research study is to investigate the impact of a Masters level mathematics teacher education course that focuses on expanding teachers' knowledge of reasoning-and-proving, in particular the extent to which the pre-service secondary teachers exhibit the following abilities over time: (1) write valid convincing arguments, (2) identify proof from non-proof solution, and (3) select or modify reasoning-and-proving tasks for their students.

Chapter two includes a review of three areas of related research and a final section on how this study is situated in and expands upon the work presented in the first three sections. The first section details students' and teachers' understandings and abilities related to reasoning-and-proving. While the focus of the investigation herein is on capturing teacher learning, exposing student struggles with reasoning-and-proving is critical since teachers need to learn students current thinking in order to support their learning. Secondly, the important role tasks play in students learning is explained. The third research area reviewed relates to teacher learning. Three professional development programs that increased teachers' knowledge of mathematics for teaching are examined. The final section of chapter two explains how the existing research described in the first three sections influenced both the design of the teacher preparation course and the research project itself.

2.1 LEARNING AND TEACHING PROOF

Opportunities for students to make original and authentic claims about mathematical statements currently are not common secondary classroom practice. Ball et al. (2002) wrote, "Proving should be part of the problem solving process with students able to mix deduction and experiment, tinker with ideas, shift between representations, conduct thought experiments, sketch and transform diagrams" (p. 912). The widely adopted Common Core State Standards for Mathematics (CCSSM, 2010) for K-12 students and the National Council of Teachers of Mathematics (NCTM) (2010) series: *Focus in High School Mathematics: Reasoning and Sense Making* expect students to be purposeful about constructing their own mathematical examples. From their examples, students will be able to identify patterns, suggest conjectures, and supply arguments to validate the truth of their claims. In response to previously proposed reform standards Thomson and Zeuli (1999) explain what teachers need to know and do to prepare students for them:

> To realize this conception of teaching, teachers need to know how to choose or design problems whose resolution will advance their students' understanding at different points along the developmental pathway toward current disciplinary knowledge, how to help students represent and express their ideas in a variety of ways, how to establish and maintain norms appropriate to a scientific or mathematical classroom community, and how to orchestrate student discourse. (p. 354)

Unpacking the recommendations expressed in the quote aligns with the first two sections of this chapter. In order for teachers to understand how to choose and design mathematical problems, they need to first understand how students think about reasoning-and-proving and their struggles to construct proofs. So the first section reviews the literature on teaching and learning proof followed by a section on choosing and designing productive tasks for increasing student knowledge of proof.

2.1.1 Secondary Students' Experiences and Understanding of Proof

In most secondary mathematics classrooms, students are not provided opportunities to make and prove conjectures, so it is not surprising that secondary students struggle to write proofs (Bell, 1976; Recio & Godino, 2001; Porteous, 1990; Senk, 1985) and are unaware of what a proof means (Chazan, 1993; Healy & Hoyles, 2000; Knuth & Sutherland, 2004). All seven studies listed collected data outside of classrooms through interviews or survey questions. In other words, the researchers did not provide an intervention to build student understanding of proof. Instead they collected data by asking questions or having students write arguments to learn if typical classroom teaching was productive (Bell, 1976; Healy & Hoyles, 2000 Recio & Godino, 2001; Senk, 1985) or to better understand how students think about justifying arguments (Chazan, 1993; Knuth & Sutherland, 2004; Lannin, 2005; Porteous, 1990). The studies, which take place outside the classroom environment, provide outcomes of what individuals thought or wrote at the moment of the survey or interview. Three other studies examined classrooms where proof tasks were implemented (Bieda, 2010; Martin, McCrone, Bower, & Dindyal, 2005; Lannin, 2005). Lannin was the researcher and instructor for 10 classes and interviewed a subset of the students outside of class. The other two studies examined classrooms and shared the challenges encountered as the teacher engaged their students in reasoning-and-proving tasks. The student verbal and written responses along with the classroom interactions portray a complete picture of what students think and how they come to understand reasoning-and-proving.

2.1.1.1 Outside Classroom Assessment: Are Examples Convincing to Students?

Many researchers suggest secondary students do not use examples to exhibit truth of a mathematics statement, but stop at examples since it is either all they understand or they have yet to gain a clear grasp of generality (Chazan, 1993; Healy & Hoyles, 2000; Lannin, 2005; Knuth & Sutherland, 2004; Porteous, 1990). Chazan (1993) learned from interviewing students that some were convinced with empirical arguments, but knew that a collection of examples was not a proof. However, Healy and Hoyles (2000) also found students when asked to write a proof overwhelmingly produced one based on examples, but when asked to identify the answer their teacher would give the best mark among a set of presented solutions, students rarely choose solutions consisting of examples alone. Additionally, Healy and Hoyles discovered that a majority of students even choose an illogical algebraic response (figure 2.1) as one their teacher would give the highest grade. The conjecture was that *when you add any 2 even numbers, your answer is always even*. The researchers concluded that these high-attaining students (aged 14-15) choose Eric's answer since it was "hard" to follow. In other words, secondary students see using examples as a sensible approach that falls short of proof, and believe teachers want to see symbolic notation even if it does not make sense to them. Both studies suggest that students exhibited a limited understanding of generality, even though they believed it was important mathematically.

Eric's Answer
Let x = any whole number
y = any whole number
x + y = z; z – x = y; z – y = x
z + z – (x + y) = x + y = 2z

So Eric says it's true.

Figure 2.1. Eric's answer

Knuth and Sutherland (2004) developed their study based on Chazan (1993), Healy and Hoyles (2000) and Porteous (1990) (discussed below) specifically focusing on the issue of generality. They argue that a limited understanding of how to generalize or what it means to generalize a situation may be an inhibitor for students' to construct proofs. Knuth and Sutherland collected data from 394 middle school students that used the *Connected Mathematics Project* (CMP) curriculum. In the study, students were presented two items. The first item asked student participants to choose between two researcher-produced solutions: an empirical argument and a proof. The solutions were derived from the following statement: *When you add any two consecutive numbers, the answer is always odd.* On this first item, 40% of the 6-8th grade students choose the solution with three examples. About 30% of the students, chose the deductive argument, which was written in words (not algebraic symbols). The remaining 30% either choose both or neither response as a proof of the statement. Based on a pilot study the previous year asking students to produce an argument for the conjecture listed above resulted in an overwhelming reliance on empirical arguments. This led the researchers to conclude that even if students could not produce a general solution, many (about 30%) recognized the need for one.

The second item asked the same 394 middle school students to write an argument to a "number trick" (*Choose a number and add 3, double the sum. Write the number down. Return to the number chosen and double it then add 6. Will these solutions always be the same?).* The question included: *Is it true for the numbers one through ten?* The small range of numbers (1 to 10) allowed for students to either write a generalization or exhaust all ten numbers to construct a proof. A follow-up question was for students to explain if the "number trick" would produce the same two numbers for any starting numbers not just one through ten. The lowest level of understanding consisted of responses in which students only produced a few examples to the "number trick" and believed that it worked for all possible numbers even those greater then ten. A second group produced a proof by exhaustion for the first part, but only chose a few examples outside the range claiming that the number trick will work for any number. A third group also produced a proof by exhaustion, but recognized the limitation for numbers greater than 10, but were unable to produce a generalization. Another group choose exhaustion and generalized for all numbers. A fifth group constructed a deductive argument from the start for all numbers, but some seemed unsure if their argument would always work. The final group constructed a general argument, but then tried a few more examples to check if their argument was secure. So this wide range of solutions suggests that some secondary students can generalize, but several issues arise. Some students exhibited deficiencies in determining the difference between using examples and proof by exhaustion. If they believe examples are enough, then they do not have a need for generality. Others see the need, but have limited skill in forming a solution for all numbers. Some are beginning to think in general ways, but struggle to understand how a general argument applies to all numbers. Overall the results make clear a wide spectrum of student thinking that teachers need to be made aware of in order to move their students along the trajectory toward generating valid arguments and understanding what generality means.

Knuth and Sutherland (2004) identified Porteous (1990) as the source for engaging students in a "number trick" to focus on the issue of generality. Porteous interviewed 50 students three times, who ranged in age from 11 to 16 years old. Unlike the Knuth and Sutherland study, students were not provided a range of numbers, so a proof by exhaustion was impossible. However, after the students answered the original "number trick", the interviewer asked if a particular number (such as 16) worked. Each interview included two general number trick statements (does it work for all numbers?) and after the student responded to the general statement, the researcher followed up by asking if a particular number was true. Porteous's rationale was that if students truly believed that the general statement was true and understood that it worked for all numbers, then a specific number like 16 would be accepted as true based on the general being true. So even if a student used a few examples and explained that it was true for all numbers, then the student would not need to check 16 if he or she believed the general case was valid. On the other hand if the student checked 16 after claiming to have proved the "number trick" for all numbers, then Porteous claimed the individual did not truly believe in their response to the general statement. An unintended result was that some students that were able to provide a logical valid argument for the general case used it to check the result of the particular case. The overwhelming majority of student responses (247 out of 290) answered the general statement with examples. Only 19 of the 247 empirical responses explained that the particular must be true since the general was true. Of the 43 (out 290) student responses that followed a logical argument, 31 used their response to the general case to answer the particular. The researcher made two overall claims from these findings. Even though students overwhelmingly rely on examples, they are not confident in empirical arguments since only 19 of 247 responses relied on a few examples to make a claim of the particular number that they did not check. On the other hand, when students that can produce a proof, they are more confident (31 out of 43) in claiming the particular must be true based on the general. Therefore, students know that deductive arguments are required to claim the truth of a mathematical statement, but most students are unsure how to move beyond just checking examples.

The final study about secondary students thinking about proof identifies an intermediate step between empirical examples and a deductive argument. Lannin (2005) was a crossover study in that the researcher taught the 25 sixth-grade students, and he interviewed four target students based on their ability levels five times each. Lannin engaged the class in solving five contextual pattern tasks over 10 class sessions. He identified generic example as a level of justification between empirical and deductive reasoning, and a successful argumentation method for students to move beyond just checking cases. A generic example proof explains general features of a situation using a particular example. The researcher uncovered that even though many students produced correct generalizations for the linear relationships, their justifications uncovered varying degrees of understanding. Some students simply guessed formulas until they found one that matched several convenient cases, and other students followed recursive methods. Both of these forms of reasoning are based on specific examples. However, successful students connected their algebraic formula with general features of a specific example: generic example. So while constructing a generalization is essential to producing a proof, examining how students construct a generalization provides more insight into their reasoning skills and trajectory toward proof. Lannin (2005) explains that knowing how students think about their generalizations will allow teachers to better support student learning of mathematically appropriate forms of justification and that some types of tasks are better suited to highlight students thinking of generality. In particular, tasks that support generic argument solutions may be a better scaffold than the leap from empirical examples to deductive arguments.

Too often students rely on examples when asked to show that a mathematical statement is true (Bell, 1976; Chazan, 1993; Healy & Hoyles, 2000; Knuth & Sutherland, 2004; Porteous, 1990; Recio & Godino, 2001). However, the consensus finding is not that students believe examples convey truth, but students are simply unsure how to construct deductive arguments (Chazan, 1993; Healy & Hoyles, 2000; Knuth & Sutherland, 2004; Porteous, 1990). Lannin suggests a generic example proof as a productive method to move students from simply checking a few examples toward deductive reasoning. Ball et al. (2002) use the term transparent proof instead of generic example and define it as "a proof of a particular case which is small enough to serve as a concrete example, yet large enough to be considered a non-specific representative of the general case" (p. 915). Ball et al. advocate introducing students to generic example proofs (transparent proof) for they are more intuitive thus more accessible to students (2002). Many of these studies suggest a progression among students from empirical examples to deductive arguments (Bell, 1976; Healy & Hoyles, 2000; Porteous, 1990; Senk, 1985) or exhaustion when possible (Knuth & Sutherland, 2004). However, informal methods of proof (generic example) appear to support students in moving away from empirical arguments since a generic argument provides students access to think in general terms while considering a specific case.

2.1.1.2 Secondary classroom handling of proof discourse

Supporting student engagement in reasoning-and-proving tasks during classroom instruction requires an environment in which students are pressed to justify their thinking. Through classroom observations of proof instruction, Harel and Rabin (2010) have identified teacher practices that are associated with an authoritative instructional view. Instructional practices that position the teacher or text as the mathematical authority in the classroom contradict calls for students to engage in problems using multiple representations. An authoritative view relies on presenting polished solutions and does not support students applying their prior knowledge to novel situations. On the other hand, the teaching of proof must stay true to the rigor of mathematics (Hanna, 1995; A. Stylianides, 2007). Students need to learn the parameters of what is an acceptable mathematical argumentation so as to not believe all explanations are valid. If teachers analyze student work and accept non-proof arguments as proof, students develop misunderstandings as to what counts. Therefore, a teacher needs to establish classroom criteria for proof to hold students accountable as they generalize patterns and justify conjectures.

Since finding tasks that promote opportunities for reasoning-and-proving are almost nonexistent in secondary curricula, Bieda (2010) choose to study classrooms using the Connected Mathematics Project (CMP) since it was identified as containing reasoning-and-proving tasks (G. Stylianides, 2009). Bieda wanted to understand whether teachers who have experience with a curriculum that includes reasoning-and-proving tasks enact them in a way that supports students in gaining an accurate view of proof. She collected data from seven middle school teacher's classrooms (3 sixth grade, 2 seventh grade, and 3 eight grade) in the same district. All seven participants taught for at least nine years, and each had taught the CMP curriculum for at least three years. Six of the seven teachers attended quarterly daylong district professional meetings on how to best enact the CMP curriculum. The researcher observed the implementation of six or seven proof-related tasks in each of the seven classrooms for a total of 43 tasks. Bieda concluded that in particular the teachers' discourse failed to support students with learning what constitutes a proof. While the students and teachers engaged in discourse, the discussions were not centered on a commonly shared view of how to determine the validity of a proposed argument. Occasionally students presented work, but the teachers rarely provided feedback. This resulted in non-proof arguments being accepted as truth. Only once did a student question another student's work. Just over half of all the proving events in the classes were justified with non-proof arguments. These results led Bieda to recommend that curricula materials provide assistance to teachers on how to provide critical feedback by recommending standards for proving.

Martin, McCrone, Bower, and Dindyal (2005) analyzed the discourse of a geometry classroom as the teacher, Mr. Drummond (Mr. D.), engaged students in proving activities. The purpose was to indentify the pedagogical moves the teacher used during instruction that supported students' understanding of valid arguments. The researchers identified revoicing student claims and requesting for student evaluation of presented arguments as two moves that appeared to support learning. Mr. D provided his students opportunities to reason and make conjectures about mathematical situations prior to justifying their arguments. For instance, students were asked to list what they noticed while examining two congruent concave pentagons. One student suggested a conjecture that the distance between two nonadjacent vertices was congruent (see figure 2.2). In particular, line segment BD is congruent to line segment NP. Mr. D suggested that the class justify this claim. After observing the teacher over a 4-month period, the researchers determined that Mr. D's facilitation of open-ended tasks was effective at providing students' opportunities to justify and construct valid arguments in a axiomatic system. However as discussed in previous research, students failed to fully grasp that a proof meant that no counterexamples are possible. Even after agreeing on a proof toward the end of the four-month observation period students still checked examples to convince themselves of the truth. Martin McCrone, Bower, and Dindyal concluded that while the discourse may have been helpful to support students in developing productive axiomatic argumentation, more is needed to help students learn that proof is not just a ritual but how truth is reached in mathematics.

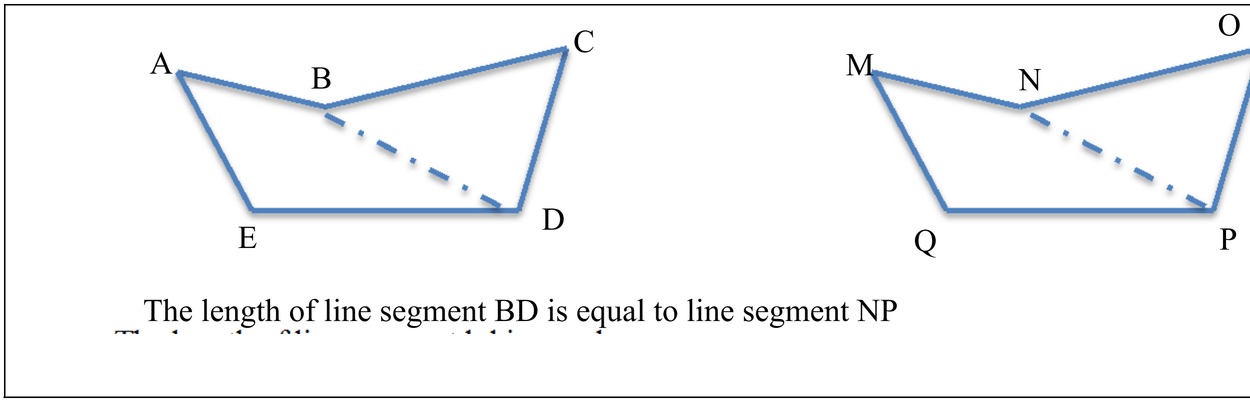


Figure 2.2. Conjecture made in Mr. D's class

Both studies (Bieda, 2010; Martin et al., 2005) identified time as a challenge to supporting students' knowledge of proof. Bieda suggested that it is possible that many of the observed class discussions may have fallen short of a valid justification because of time restrictions. Additionally, Martin et al. identified time on occasions when Mr. D took over the thinking of students. They noted that Mr. D presented students proofs when the end of class was approaching. Knuth (2002b) found that teachers also indicated that time would be a reason for not enacting proof tasks in their classrooms. This obstacle deserves further consideration since teachers identify it as a reason for ignoring proof or limiting students to fully explore it, and if justification is truncated as a result of time, it can confuse students about what are acceptable arguments.

Another interesting observation in both classroom studies was the absence of any mention of the use of classroom criterion for what counts as proof. Bieda (2010) identified A. Stylianides (2007) criteria as her method for analyzing the classroom argumentations, but did not make it clear whether any of the classroom teachers constructed a list of proof characteristics with their students. Additionally, Mr. D was noted as using productive classroom discourse, but there was no mention of a criterion that students in the class could use to hold one another accountable. A. Stylianides (2007) explains that an explicit definition of proof serves two important functions for instruction: 1) connecting the classroom with the mathematics community where the teacher is the link, and 2) being explicit as to what counts and using the list to validate solutions. When classroom communities do not create a shared criterion for judging proofs, it is possible for non-proof arguments to be accepted (A.J. Stylianides, 2007), as was the case in more than half of the implementations in the Bieda study.

Most of the research on proof with secondary students focuses on students' lack of understanding on two levels: 1) *students possess a limited understanding of how to construct a valid argument*; and 2) *students possess an insecure knowledge of what a proof means in mathematics*. For instance, some students can construct a valid argument, but are uncertain when questioned about the arguments generality. Opportunities for students to produce generic examples have shown to be a productive scaffold in moving students between empirical and deductive arguments. Classroom discourse plays a critical role with developing students understanding and without an explicit agreed upon list of what counts as proof, invalid responses may be accepted. Accepting non-proof arguments as proof is counter productive to developing students thinking.

2.1.2 College Students' and Teachers' Experiences with Learning to Prove

Upon graduation from high school and prior to teaching secondary mathematics, the final opportunity to learn to prove statements occurs in undergraduate mathematics courses since most teacher certification programs do not offer specific courses on reasoning-and-proving. In high school, students typically only study proof in an axiomatic Euclidian geometry course, so their transition to proof at the university level is unexpected and abrupt (Moore, 1994). Students enter novel fields of mathematics such as abstract and linear algebra and are exposed to new terms in these domains. Instead of manipulating equations or following route procedures as they did in most high school courses, students are called upon to make sense of definitions in new mathematical domains and apply them to form valid arguments. Needless to say, undergraduate mathematics majors struggle to understand novel concepts, so applying them to proof writing is a difficult undertaking (Edwards & Ward, 2004; Moore, 1994).

In interviewing first year undergraduate students, Solomon (2006) investigated their previous experience and current understanding of proof. In particular, she interviewed 12 students to gain their insight on proof construction and their role as students in the formulation of arguments. Similar to Harel and Rabin's (2010) discussion of authoritative practices, the participants in Solomon's study proclaimed to be outsiders in the negotiation of conjectures and exploration of patterns. Instead they labeled the professor as the authority in the construction of proofs and indicate that students were only asked to reproduce arguments previously presented in class or were shown proofs as side activities. The proofs were presented to tell the students why, opposed to the class constructing proofs to understand why a method works or why a conjecture is true. For instance, one student was quoted as saying: "I'm told 'so and so and so and so is this' then I won't go and read and try and understand why. I just remember the result… I think they just do it so they don't get criticism of just throwing it at you" (p. 387). Consequently, these undergraduate students did not view proof as playing an integral part of mathematics. Solomon argues that transforming student views of proof is only possible through changing pedagogical practices away from computational and individual result driven instruction, toward whole class construction and communication. The implementation of proof tasks heavily influences students' perceptions and beliefs of proof (Solomon, 2006; Harel & Rabin, 2010).

It is possible for students to make sense of concepts if they are provided time to investigate worthwhile tasks. However, they need to learn reasoning skills to understand how to make sense of new concepts. Dahlberb and Housman (1997) conducted cognitive interviews with 11 third and fourth year mathematics students at a small liberal arts college. Most of the students were considering careers as secondary mathematics teachers. Students were provided a fictional definition related to a concept they were asked to explore with specific directions. Then the interviewees were asked to verify if six suggested statements met the concept definition. The final part of the interview was list of four conjectures the students were asked to prove (as shown in figure 2.3). This study was based on the ideas of concept image (definition) and concept usage (application) initiated by Tall and Vinner (1981). Students that made use of generating examples and representations were most successful at identifying correct conjectures. Students that used other strategies such as memorization to understand the concepts were less successful and usually guessed (incorrectly) without much justification when asked to identify true conjectures. Additionally, students who utilized memorization to learn the concept relied heavily on the interviewer to determine correctness of their answers. Providing students opportunities to understand the value of generating their own examples or constructing diagrams to understand the context of a problem better prepared students to make conjectures, and allowed them to become more reliant and confident with their mathematical ability. While this study was based on interviews, Dahlberg and Housman recommend that classroom instruction should promote students to generate their own examples and connect multiple representations to support students in developing proof arguments.

Definition	**Instructions**	**Verification page (Determine and justify**	**Conjecture page**

		which are fine)	
A function is called fine if it has a root (zero) at each integer.	a. Give an example of a fine function and explain why it is a fine function. b. Give an example of a function, which is not fine and explain why it is not fine. c. In your own words and/or pictures, explain what a fine function is.	f(x) = sin(πx) f(x) = x2 – x f(x) = 0 f(x) = 0 if x is rational; 1 if x is irrational f(x) = tan((π/2)x) a graph	No polynomial is a fine function. All trigonometric functions are fine. All fine functions are periodic. The product of a fine function and any other function is a fine function.

Figure 2.3. Contents of Dahlberg and Housman interview

Smith (2006) interviewed five students in two different introductory number theory courses at a large state university. Two students were in enrolled in the introductory number theory class taught in a lecture format, and the students' role was to passively follow the professor as he solved problems. The other three students were enrolled in the same number theory course, but the professor engaged students in a problem-based format where students were expected to actively engage in solving problems in class. The students were interviewed twice during the semester course. The interviewer asked students of their views of proof and the role it played in mathematics and to "think aloud" as they proved two number theory statements. The two interviews contained both parts (questions, solving problems), but the second interview had an additional section where the students were asked to validate arguments. All five participants reviewed four solutions in the second interview that were adapted answers the students themselves constructed during the first interview.

The students in the traditional course described proof with a strong focus on structure and form, while students in the problem-based course valued meaning. Consequently, the students in the traditional course viewed proof as an algorithmic process, and the others focused on making sense of the concepts and writing down what they knew about the problem as they crafted their argument. Needless to say, the traditional course participants were reluctant to generate their own examples to gain insight into problem context. With regard to analyzing arguments, the students in problem-based class again applied an understanding versus judging based on form. For instance, the two students in the traditional course praised a solution that used a proof by contradiction method even though there was an obvious gap in the argument. The results of the study suggest a relationship between classroom instruction and students' practices with solving proof tasks and analyzing arguments.

Traditionally presenting undergraduate students with proofs is problematic for a few reasons. It distances them from understanding the purpose of proof in mathematics, and denies them access and belief that they can construct a proof on their own. A presented proof becomes an object for students to memorize (Knuth, 2002b). Providing college students an environment in which they can reason through example generation shows promising results toward improving their ability to write proofs as was shown with secondary students. It is well documented that university students struggle to write proofs in university courses mostly because the students are unfamiliar with the content and forms of reasoning (Weber, 2001). More specifically, undergraduate math majors cannot write proofs because they do not know the definitions of the terms in the problem (Edwards & Ward, 2004; Moore, 1994). However, even when the concepts are at the high school level students still struggle. Selden and Selden (2003) interviewed eight undergraduate math and math education majors and found that only two were able to write a proof for the this statement: *For any positive integer n, if n^2 is a multiple of 3, then n is a multiple of 3.* So even college students majoring in mathematics need classrooms where they are supported to reason so they can make sense of new ideas, organize valid arguments, learn what constitutes a proof, and understand the role proof plays in mathematics.

As previously mentioned, most practicing teachers are not exposed to additional courses or learning experiences to learn proof beyond undergraduate courses. So it should not be surprising that secondary teachers believed it is their responsibility to present a proof to their students since this was what they experienced and expected as students. And since college students struggle to produce proofs, it should also not be surprising that teachers believe writing them is difficult. Therefore, expecting high school students to prove statements beyond traditional two-column geometry proofs is unlikely. Knuth (2002a; 2002b) explored these issues. As with secondary and university students, he reported that experienced teachers also demonstrated a limited view of proof. Secondary teachers misunderstand the role proof plays in mathematics, do not believe all students should be exposed to proof tasks (Knuth 2002b), and lack an accurate criterion for evaluating student arguments (Knuth 2002a; Selden & Selden, 2003). In others words, teachers not only need to learn how to support students learning, but they also have to learn for themselves about the role of proof in mathematics and need additional support with constructing logical valid arguments. In particular they saw it as a separate topic of study rather than a way to makes sense of content. Teachers conveyed a limited understanding of which arguments count as proof. In addition to limiting proof to special topics, they also believed it should be reserved only for the highest achieving students and not an activity in which all students could or should participate. Restricting access and opportunity to reason-and-prove contradicts what mathematics educators and standards have suggested as its role in schools (Ball et al. 2002; Hanna, 1995; NCTM, 2000; CCSS, 2010). Providing teachers an opportunity to engage in tasks that provide access for more students across a variety of contexts could help them to realize how proof can be accessible to all students and applicable in beyond special occasions.

2.2 REASONING-AND-PROVING TASKS

Instructional tasks provide students opportunities to learn concepts, and shape students thinking about the subject in general (Doyle, 1983; NCTM, 1991). In other words, if teachers only engage students in tasks that require them to follow a provided procedure, then students will only improve their ability to carry out procedures and believe that to study mathematics and gain competence is to perform procedures. Moreover, if students are asked to engage in classroom tasks that promote reasoning, then students will not only see mathematics as a creative process, but will also become proficient. "Worthwhile tasks" not only address the topics in the grade level curriculum, but also provide access to a diverse group of learners, allow for more than one correct answer, and stimulate students to interact with one another as they reason (NCTM, 1991).

Teachers' ability to choose or design reasoning-and-proving problems is especially important since many secondary curricula contain a limited supply of tasks requiring these processes (Johnson, Thompson, & Senk, 2010). G. Stylianides (2008) designed the reasoning-and-proving framework as a research tool, and since explained how teachers could use it as a trajectory for scaffolding students thinking toward valid arguments (G. Stylianides, 2010). A mathematically acceptable criterion for judging the validity of proofs should be constructed in classrooms so that students can hold each other accountable (A. Stylianides, 2007). Specific reasoning-and-proving discourse would address both moving students along the activities in the framework and contrasting whole class presented arguments against the develop criterion. Preparing teachers to choose worthwhile tasks and support their students with understanding and constructing proofs in this way is half of the challenge.

Doyle (1988) introduced the concepts "cognitive level" and "academic demands" of a task. He described tasks that prompt students to recognize or memorize information such as multiplication facts or provide the name of a geometric shape as a low cognitive level tasks. A high cognitive level task promotes problem solving. Mathematics standards (NCTM, 1989, 2000; CCSS, 2010) promote both procedural and conceptual understanding, but the majority of tasks in secondary curricular materials do not provide opportunities for students to develop reasoning skills (Hanna, 1995; Johnson, Thomson, & Senk, 2010).

The researchers on the QUASAR project also noticed the importance of tasks and their affect on student learning (Stein, Grover, & Hennigsen, 1996). Stein and Smith (1998) expanded upon Doyle's (1988) work with cognitive levels, later articulated in the Task Analysis Guide (TAG). Similar to Doyle's levels, the TAG is divided into low and high cognitive levels of demand. However, Stein and Smith specified the TAG for specifically analyzing mathematics tasks as opposed to Doyle's subject neutral descriptions. The two lower level cognitive categories are *Memorization* and *Procedures Without Connections*. *Procedures with Connections* and *Doing Mathematics* are the titles of the high cognitive demand levels. Each of the four levels possesses distinctive qualities.

Low-level tasks lack a press for conceptual understanding or justification, and focus on producing one correct answer (Stein, Smith, Henningsen, & Silver, 2000). Memorization tasks involve students recalling previously learned concepts in which no procedure is needed or the amount of time allotted for the task restrains the possibility of following one. Asking students to list a definition or theorem is an example of a memorization task. Procedures without connections tasks require a procedure, but suggest following a method without explaining why it works or how to relate multiple representations. An example could be to ask students to write an equation from coordinate points in a given x-y table. A scripted rehearsed procedure is the expected solution path such as: the y-intercept is substituted for b in the slope-intercept form of a linear equation (y = mx + b) and m is replaced with the fraction comprised of the change in the y values in the numerator and difference between the x values in the denominator. While this is a popular procedure taught in pre-algebra and algebra classrooms, it does not explain why or in which situations this process works. A procedure without connections might ask students to explain their procedure, but the explanation only retells the steps without attention to mathematical understanding. Additionally, connections to other representations such as a graph, which could lead to conceptual understanding is ignored. Overall low-level tasks require limited thinking on behalf of the student, but can be used to improve speed and precision with routine problems (Stein et al., 2000).

High-level cognitively demanding tasks are intended to build a deep understanding of particular concepts and gain a greater sense of what mathematics is in general (Doyle, 1988; NCTM, 1991; Stein et al., 2000). Procedures with connections tasks allow for students to choose a solution path based on their prior knowledge or to draw comparisons across multiple representations or methods (Stein et al., 2000). While a process is used, it is not followed without thinking through the problems context or underlying meaning. An example of a procedures-with-connections task would be the following problem:

Tim has $1,000 and places it in a bank, which earns a simple 5.5% annual interest rate. Ginny also has $1,000 saved and finds a bank that offers a 5% compounded quarterly interest rate. Ginny tells Tim to move his money into her bank because in 6 years when they graduate high school she will have more money than him. Tim says that is impossible my bank provides a greater interest rate. Make a graph of the first six years of Tim and Ginny's money and explain who has the better savings plan.

The Tim and Ginny bank task is procedures with connections since it expects students to follow the procedure to find simple and quarterly compound interest, and tells students to make a graph. However, as students follow the procedures they will notice the differences between the two savings plans and more generally simple versus compounded interest. The problem is not simply asking for how much each child has in the bank in six year. Doing mathematics tasks are the highest level and require students to investigate novel problems and at times, multiple solutions. The ambiguity of these doing math problems may cause students to become frustrated for they require sustained attention to try various methods and simultaneously keep track and organize successful and failed attempts to recognize patterns. An example of a doing mathematics task is:

The Glee Club wants to order shirts for their 5K fundraising event. Last year 250 people ran in the race. This year they expect many more runners based on the club's Facebook page. Two shirt companies expressed interest in providing support through offering special discounted prices for the fundraiser. Tina's T-shirt shop will charge $40 to create the 5K running logo and $6 for the first 200 shirts and $4 per additional shirt beyond 200. Stevie's Shirts offers $100 to create the shirt logo and $5 per shirt. Make an argument to convince the Glee Club in support of one of the shirt companies.

The t-shirt buying task is at the doing mathematics level since there is no correct answer. Students will need to develop a contextual argument based on their estimates for how many runners they expect at the race. They could choose to solve the problem in several different ways. Extensive practice with high-level tasks will improve students' ability to solve problems and reason in a variety of contexts. As the standards call for more student understanding (NCTM, 1989, 2000; CCSS, 2010), these cognitively demanding tasks are gaining extensive attention in mathematics research and professional develop since they are scarce in curriculum materials (Hanna, 1995; Johnson, Thomson, & Senk, 2010) and are difficult for teachers to implement (Henningsen & Stein, 1997; Stein, Grover, & Henningsen, 1996).

In addition to the work of Stein and colleagues on tasks in general, Stylianides (2008, 2010) proposed a framework for looking at reasoning-and-proving in particular, which includes three components (Figure 2.4). The term, reasoning-and-proving is hyphenated to include the full range of activities associated with scaffolding students' thinking with constructing proofs. The mathematical component includes two sections generalization and argumentation. Explicitly requiring students to first examine cases to find a pattern provides students access to begin thinking about a mathematical situation. Once students observe regularity within a pattern that they constructed, they are better prepared to suggest a conjecture. Both looking for patterns and making conjectures contribute to developing a generalization. Arguments are simply non-proofs or proofs, with two types in each category. While empirical arguments are not proofs, starting by generating examples helps students to make sense of the mathematics, which can lead to developing a proof (Dahlberb & Housman, 1997; Lannin, 2005; Smith, 2006). A rationale, also not a proof, is not example based, but the solution makes logical leaps or includes statements that have yet to be accepted by the mathematical classroom community (Stylianides, 2008). All proof tasks do not need to include the full range of activities, but initially supplying students with tasks that first allows them to look for patterns provides a scaffold toward a generalization and proof. Teachers should choose tasks that explicitly call for the generalization activities, so that students come to realize how the activities in the framework are helpful in producing a proof. The hope then would be that students would look for patterns and make conjectures even when tasks are more open-ended.

!

<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 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?</td></tr>
</table>

Figure 2.4. Reasoning-and-Proving Framework adapted from G. Stylianides (2010)

The learner component focuses on the students’ conception of the four different arguments. The teacher questions the learner to try and uncover misunderstandings about the nature of a proof. Porteous (1990) tested learners’ understanding of generality when he asked the students to determine if the number 16 would work after students already were expected to write a proof. How the students answered the question, provided the researcher with insight into the learners thinking of proof. Knuth and Sutherland (2004) also questioned students thinking of proof and noticed some students continued to examine cases after they claimed to have generated a valid argument. In other words, known student misconceptions about reasoning-and-proving are turned into questions to press students thinking toward a broader understanding of proof beyond just constructing valid arguments.

The pedagogical component builds on both the learner and mathematical components. Based on what the teacher learns from engaging with his or her students, the teacher will need to make decisions that connect the students' current thinking and to the more conventional understanding of the broader mathematical community. Here the teacher is positioned to choose tasks or choose specific questions to bridge students' knowledge. For instance, students who seem satisfied with empirical arguments could change their thinking if they were asked to solve a task where the initial pattern fails after the first few examples. Engaging in such a task could provide the students with a need for something more than checking a few cases and becoming convinced of its truth (G. Stylianides & A. Stylianides, 2009). Once students believe that a few examples is not enough, Lannin (2004) suggested providing students with tasks that provide an opportunity for them to generate generic arguments since such tasks are useful at bridging students thinking from empirical to deductive thinking. A teacher is attending to the pedagogical component of proof as he or she specifically chooses tasks that foster students' growth along the trajectory from non-proof to proofs and illuminates their knowledge of each of the arguments.

Proof tasks do span all four levels (as shown in Figure 2.5). The memorization task in the top left corner of the figure (2.5) only asks students to fill in blanks of an almost complete two-column proof. Students are expected to recall reasons or statements to complete a very structured and rigid argument. Memorization proof tasks do not engage students in any of the activities listed in the reasoning-and-proving framework. Related to procedures without connections, most pre-service secondary mathematics teachers do not understand how the multiple steps used in the process of mathematical induction proves conjectures true (G. Stylianides, A. Stylianides, & Philippou, 2007). Using a procedure to produce a solution without reasoning or a complete understanding of how or why the procedure works is the essence of the procedures without connections category. Therefore, low-level tasks call for the completion of a proof, and neglect the opportunity for students to reason and justify their thinking.

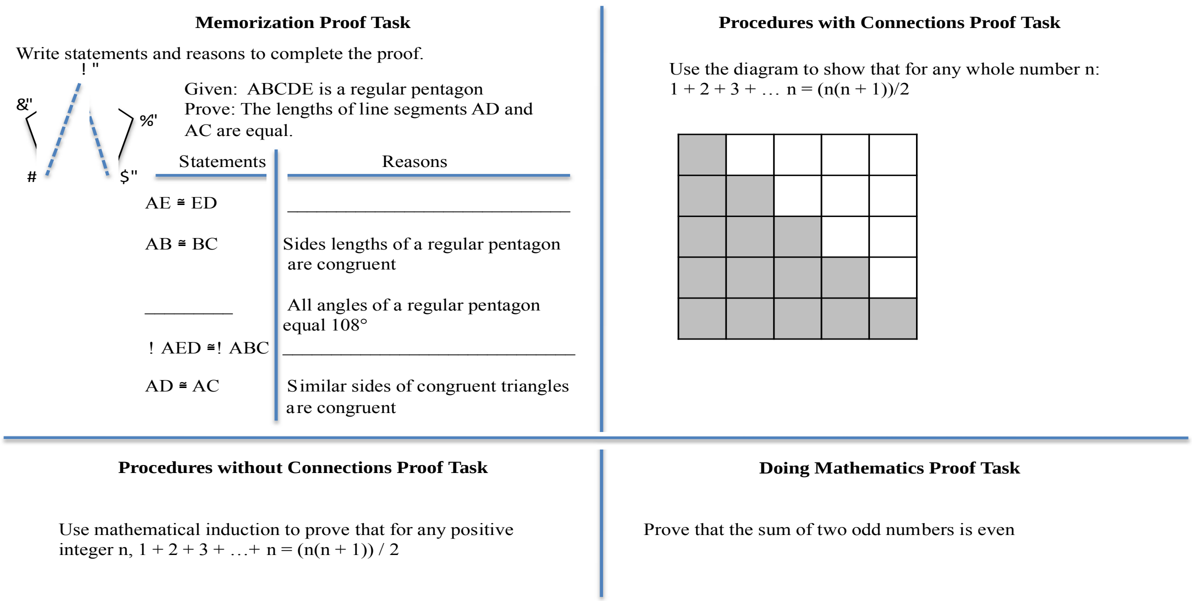


Figure 2.5 Cognitive levels of proof tasks

A *procedures with connections* proof task is an example of making an explicit connection between an equation and diagram. There are multiple ways to prove the conjecture, but connecting the two representations would help students develop understanding as to the truth of the statement. A student could start with examples such as a 3x3 square and relate that to the sum of the first three counting numbers and build a pattern of more examples. Starting with examples can scaffold students thinking toward constructing a generic example, which is helpful in bridging thinking from empirical to deductive arguments (Ball et al., 2002; Lannin, 2005; G. Stylianides, 2008). A doing mathematics proof task is not explicit about how to start the problem and could be frustrating for some students. These tasks are useful to assess students understanding of the usefulness of reasoning activities to generate a proof. In other words, would students know to generate examples on their own without the problem explicitly requesting them? Additionally, it could be used to learn which students are able to generalize or construct a demonstration. High-level cognitively demanding proof tasks either explicitly or implicitly engage students in the full range of reasoning-and-proving activities to develop justification for proposed conjectures.

While there is an abundant amount of research, which points to secondary students' inability to write proofs (e.g. Bell, 1976, Chazan, 1993; Healy & Hoyles, 2000; Lannin, 2005; Knuth & Sutherland, 2004, Porteous, 1990, Senk, 1985), research is scarce on how to support students learning in the domain. Knuth and Sutherland argue that "If more students are to develop their understanding of generality – and of proving more specifically – then they must be given opportunities to engage in activities which highlight important ideas about proving" (2004, p. 7). High-level mathematics tasks provide students an opportunity to develop understanding. Furthermore, G. Stylianides's (2010) reasoning-and-proving framework provides a full range of activities that provide students access to proof. Therefore, engaging more students more often in high-level reasoning-and-proving tasks along with serious considerations of both the learner and pedagogical components of the framework is a promising path with supporting students in exceeding the recommendations detailed in the policy documents (NCTM, 2000, 2010; CCSS, 2010). However, a serious challenge is to prepare teachers to select or design reasoning-and-proving tasks since many secondary curricula contain a limited supply of tasks requiring these processes (Johnson, Thompson, & Senk, 2010).

The following section will provide a theoretical rational for professional development. In other words, what do teachers need to know? Three examples of professional development are provided for empirical evidence. The results of these cases of teacher learning will then be explained in how they are applied to the learning situation for the participants in this study.

2.3 THEORETICAL PERSPECTIVE FOR PREPARING TEACHERS TO IMPLEMENT R&P TASKS

Prior to the publication of the earliest standards document, Shulman proposed the existence of a "knowledge base for teaching" (1987, p.4). While Ball, Lubienski, and Mewborn (2001) credit Shulman for introducing a knowledge base for teaching, they point to the importance as common sense. While the theory of a certain knowledge base for teaching seems obvious, actually identifying what teachers should know and how they might come to know content and skills to successfully engage students in learning mathematics is not obvious (Ball, Lubienski, & Mewborn, 2001; Shulman, 1987).

Shulman (1987) listed seven types of teacher knowledge and identified pedagogical content knowledge as the one of special interest. Over two decades later, Ball, Thames, and Phelps ask, "What have we learned and what do we yet need to understand [about pedagogical content knowledge]" (2008, p. 392)? Their conclusion was that the research field has not made much progress on reaching Shulman's vision of building a theoretical framework of a knowledge base for teaching.

Following a related but more practical perspective, Doyle (1983) explained that the tasks teachers provide to students in the classroom strongly influence students' thinking about the content. The QUASAR[3] project team expanded upon Doyle's concept of academic task from both a practical and theoretical perspective. The Mathematical Tasks Framework (MTF) (as shown in figure 2.6) was created and applied to analyze classroom instruction (e.g. Boston & Smith, 2010; Stein, Grover, & Henningsen, 1996; Henningsen & Stein, 1997) and was used as a conceptual instructional tool in developing professional development materials (e.g. Stein, Smith, Henningsen, & Silver, 2009).

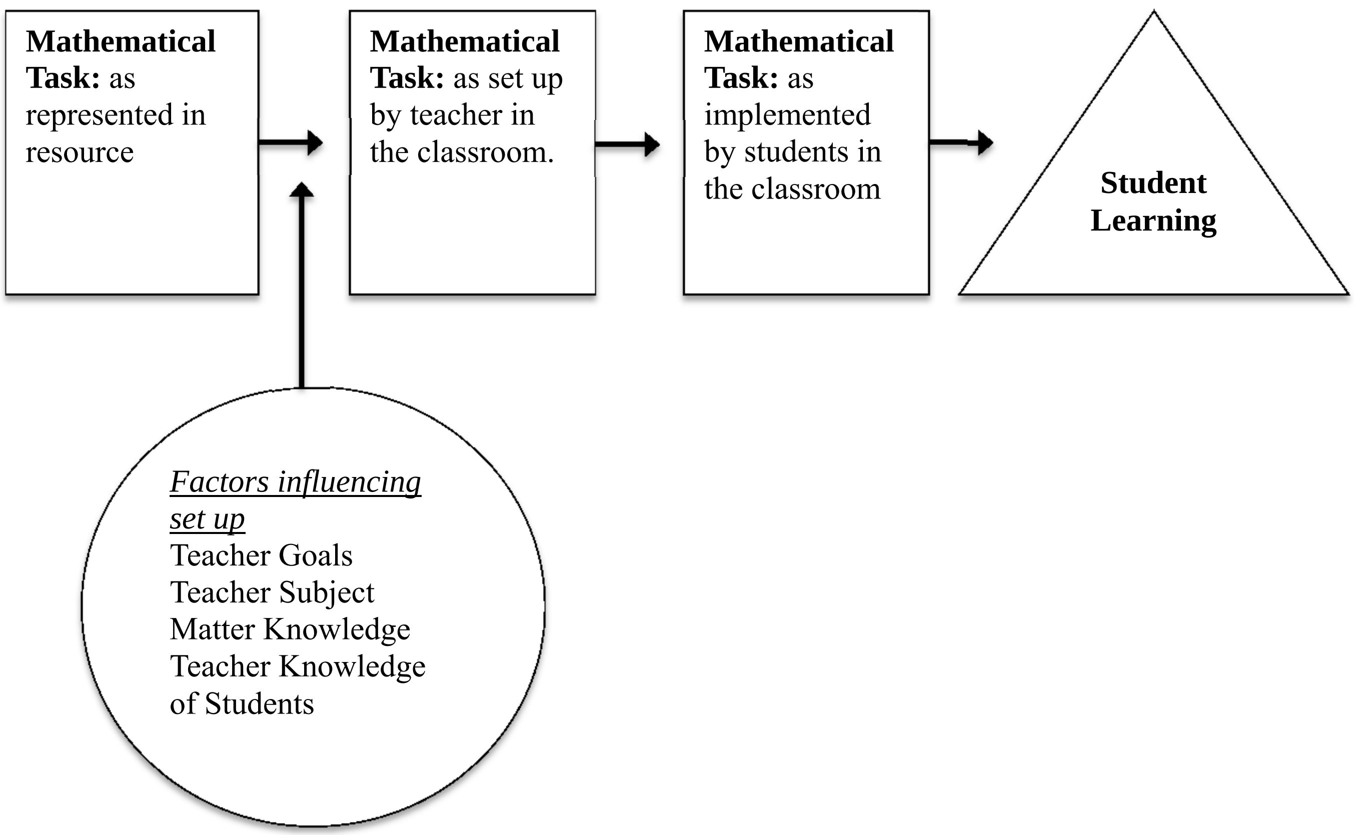


Figure 2.6. Mathematical tasks framework adapted from (Stein & Lane, 1996)

[3] The QUASAR (Quantitative Understanding: Amplifying Student Achievement and Reasoning) Project was a national reform project aimed at assisting schools in economically disadvantaged communities to develop middle school mathematics programs that emphasized thinking, reasoning, and problem-solving (Silver & Stein, 1996).

The Mathematical Tasks Framework (MTF) is a conceptual perspective that can be used to think about how instructional tasks unfold prior to and during classroom instruction (Stein et al., 2009). The three rectangles represent the phases through which a task passes as it moves from selection to implementation. The theoretical hypothesis is that choosing high-level tasks during the selection process (first rectangle), and maintaining the high level demands of the task during the subsequent two phases (second and third rectangles) results in student learning (the triangle).

The circle between phases 1 and 2 in Figure 2.6 includes the factors that influence the task set up. In addition to the intended learning goals, knowledge needed for teaching is listed. Similar to Shulman's (1987) identification of knowledge needed for teaching, Stein and Lane (1996) also recognized the importance that teacher knowledge contributes to their ability to implement a mathematics task. In other words, knowledge needed for teaching is embedded in the MTF where a teacher's understanding of the content and students thinking influences their instructional decisions.

Supporting teachers to select and enact reasoning-and-proving tasks (the three phases of the framework) is the basis of the course embedded within this design experiment. For instance, solving tasks was intended to build content knowledge to make better instructional decisions. Analyzing narrative cases allowed for the participants to reflect on their own instruction and identify factors that support learning at each of the three phases. However, this study is focused on the first phase including the knowledge of reasoning-and-proving and how student think about the domain since they are factors that influence the task set up. Pilot data from the initial implementation of the course materials showed that teachers struggled to identify high-level reasoning-and-proving tasks. Starting with low-level tasks rarely provides students with opportunities to reason during implementation (Stein et al., 1996). So knowing how to produce a proof or identify valid arguments alone was not enough support for teachers to select of modify high-level reasoning-and-proving tasks. Explicit instruction was provided to the participants to identify a task in addition to solving tasks and analyzing student work with respect to proof.

The last section of this chapter will explain more about how this study will expand on the first phase of the MTF along with the knowledge needed to set-up a reasoning-and-proving task.

2.4 THREE EXAMPLES OF PRODUCTIVE PROFESSIONAL DEVELOPMENT THAT EXPANDED TEACHER KNOWLEDGE

The previous sections focused on what students understand about reasoning-and-proving, how classrooms can support learning, the types of tasks that promote reasoning-and-proving, and a theoretical model for task unfolding. Through reviewing the content and context of productive professional development programs, ones that expand teachers' knowledge and change their practice, this section will glean the properties from the programs that make them successful. The previous section helps the mathematics education community understand what teachers need to know, but the question addressed in this section is *how* teachers might best learn the knowledge for teaching reasoning-and-proving.

The focus on supporting teachers' efforts to improve student-learning outcomes is a new area of study when considering the history of education in this country. Prior to the latest reform movement, teaching was described as an autonomic occupation, not a learned profession (Shulman, 1987). However, research has provided evidence that teacher knowledge and classroom practice can change given a sustained and focused professional development program

(Boston & Smith, 2009; Carpenter, Fennema, Peterson, Chiang, & Loef, 1989; Simon & Schifter, 1991). While policymakers are eager to identify the relationship between professional development and improved student-learning outcomes (Guskey & Yoon, 2009), finding how teachers internalize concepts as learners and enact them in their classroom are essential intermediate steps. Scher and O'Reilly (2009) designed a theoretical model where they identify three stages of professional development. The first step includes expanding teachers' content and pedagogical knowledge and dispositions. The middle phase of growth is focused on teachers' change in instructional practices to match their expanded knowledge and beliefs. The final outcome to professional development is increased student achievement and change in student attitudes. Their theoretical model implies that student improvements are dependent on teacher growth, or without a change in teacher knowledge, student achievement will not improve. They do not suggest the teacher learning as a linear model. Instead teachers would engage in activities to increase their knowledge content and practice outside of the classroom. This model aligns with the productive professional development programs that will be explained.

The content of the professional development learning programs that have shown a positive change on instructional practices and teacher knowledge address three critical areas: 1) building content knowledge, 2) students' thinking about and learning of the subject, and 3) pedagogical skills (Borko, 2004; Guskey & Yoon, 2009; Scher & O'Reilly, 2009; Thompson & Zeuli, 1999). In order to learn these three constructs for teaching, the professional development curriculum materials should be situated in the everyday practice of teaching (Ball & Cohen, 1999; Putnam & Borko, 2000; Smith, 2000). For example, narrative cases of teacher instruction have shown to improve pedagogical knowledge (Barnett, 1993). Others authentic activities of practice include but are not limited to videos, analyzing student work, solving math problems, and others related to planning instruction such as anticipating student thinking. For learning these materials to occur, a level of disequilibrium must occur (Ball & Cohen, 1999). Teachers need to experience a cognitive conflict between their current thinking and or beliefs of the content, pedagogy, or students with their engagement with materials as learners. In other words, if teachers interpret their experiences as consistent with their current practices or understanding of the content, then change is unwarranted in the mind of the teacher. To change teachers' knowledge and beliefs about content, pedagogy, and student thinking, professional development programs need to engage teachers in tasks and pedagogy that are grounded in their everyday practice and cause a sense of disequilibrium.

The following three sections will examine the contexts, content, and effects on teacher learning from three separate professional develop programs: Educational Leaders in Mathematics (ELM), Cognitively Guided Instruction (CGI), and Enhancing Secondary Mathematics Teacher Preparation (ESP). These three programs were chosen based on the aforementioned characteristics of situating the professional learning in everyday teacher activities and the premise that change in thinking and practice occurs when the learner enters a conflict between their current habitual understanding and an unfamiliar yet rationale situation. Additionally, CGI and ELM are identified as exemplars programs in mathematics education (Borko, 2004; Punam & Borko, 2004; Thompson & Zeuli, 1999) and ESP is a more recent study of secondary mathematics teachers that references both of the other two studies and incorporates ideas learned from another well-respected and referenced program, QUASAR. This current study draws on the ELM, CGI, and ESP professional development programs in practice and research methods to identify changes in teacher knowledge and instruction.

2.4.1 Educational Leaders in Mathematics (ELM)

The Educational Leaders in Mathematics (ELM) was created to prepare in-service teachers for the demands of the *Curriculum and Evaluation Standards for School Mathematics* (NCTM, 1989) as both a research and instructional program (Simon & Schifter, 1991). The program supported teachers' growth in understanding the standards, and studied the effects of learning on both the participants and their students. ELM follows a constructivist perspective of learning. In the classroom, the vision is for students to construct their own meaning of mathematical situations. A mental dissidence occurs between the learners' current understanding and realizations that result from engagement on a task. The disequilibrium causes the individual to modify his or her knowledge by negotiating prior thinking with new experiences. Additionally, the classroom community of learners' develops a shared understanding based on individual contributions. Three founding guidelines provide structure for the ELM program (Simon & Schifter, 1991 p. 312):

> *1) Teachers must be encouraged to examine the nature of mathematics and the process of learning mathematics as a basis for deciding how to teach mathematics,*
>
> 2) *Teachers' learning can be viewed in much the same way as mathematics students' learning, and*
>
> 3) *Provide follow-up supervision and support.*

The ELM professional development program prepared teachers to support their students in constructing mathematical knowledge through four stages of development. The four stages consisted of a two-week summer program (stage 1), follow-up though classroom support (stage 2), planning sessions (stage 3), and lead local professional development (stage 4).

The first ELM stage of professional development for 7-12th grade teachers included professional learning sessions. Stage one was a two-week long summer program, which included three courses. Course one provided an opportunity for teachers to learn math and discuss solutions followed by reflecting on the learning experiences. In this course teachers also analyzed and discussed videotaped interviews to study students' thinking and planned lessons in grade level groups for future implementation around critical concepts. The second course involved working with computer software programs as a tool to study shapes and, in particular, to make and verify conjectures. During the final course teachers learned tennis for a week and jazz dance instruction for the second week. The three-course two-week summer program provided teachers opportunities to explore mathematics topics in new ways and understand what it is like to learn unfamiliar activities such as dance or tennis.

Following the summer program, ELM project staff provided classroom support and additional learning opportunities to foster student conceptual understanding, which was considered the second stage. In the following school year after the summer program, an ELM staff member attended each teacher's classroom for one period once per week. The weekly in class observations included a thirty-minute instructional follow-up. On occasion the ELM facilitator taught a portion of a lesson. In addition to the weekly observations, the teachers attended four workshops during the school year to discuss implementation and revisit the summer learning activities: solving tasks, analyzing student understandings, and planning lessons. The third stage was for the teachers that applied after participating in stage 2. While teachers continued to reason about mathematical content and examine how students enter and make sense of the same concepts, the focus shifted toward planning instructional lessons. Teachers examined published curriculum and were asked to adapt the written text problems and plan lessons to incorporate their new understanding of teaching and learning. Another piece of this stage was to prepare the teachers for new roles as instructional leaders. The final stage (four) was an effort to scale up the ELM program. The teachers took on even greater leadership roles as they facilitated ELM workshops. So the teachers started as learners in the initial stage, taught lessons in their classrooms to refine their pedagogy in stage two, gained more experience planning lessons in the third ELM stage, and finally were asked to facilitate workshops with their colleagues.

The ELM staff members collected teacher writings and conducted interviews to study the programs effect on teacher knowledge and beliefs. The researchers identified several themes that emerged upon reviewing teacher written reflections about their experiences at the conclusion of the first stage. Teachers' written reflections conveyed new insights into thinking about how students learn mathematics, and their role as teachers in supporting students' growth. The features of the professional learning environment identified as positive contributors to their new thinking about teaching included: engaging participants in small group work, the modeling teaching with thought provoking questions, providing time for groups to explain their current understanding of a problem, and observing how their colleagues solved tasks. One teacher indicated that stage one created "disequilibrium" between his past teaching and learning experiences and this new opportunity. Specifically addressing how his role as a learner of mathematics shifted from mindlessly applying procedures to solve problems to actively thinking of his solution paths and reflecting on his chosen methods.

During the follow-up year, teachers completed questionnaires to comment on how the ideas learned in the summer session were working during the school year. They wrote that they listen more to what students were saying as part of integrating the ELM strategies. Focus shifted from getting the right answer to actively involving all students in building conceptual understanding. While the first year was not considered ideal, some teachers mentioned that this was just the first step in trying to improve instruction. In general the self-reports pointed toward shifts in their beliefs about mathematics and how to teach it, and suggested that they had gained new knowledge about their role in supporting students in an organized productive learning environment.

A total of 56 teachers were interviewed at the end of stage two and 15 teachers that continued until the final stage were interviewed a second time. The interviews were assessed on two 5-point scales (0, III, IVa, IVb, V): ACMI: *Assessment of Constructivism in Mathematics Instruction* and LoU: *Levels of Use.* Both scales range from lowest level (not using a particular strategy: LoU or no use of constructivist epistemology: ACMI) to the highest is (collaborating or assisting colleagues with implementing the programs practices). The ELM staff identified nine different instructional strategies that the interview raters tagged and rated the teacher's level of use (LoU). For instance, asking non-leading questions and using non-routine problems were two of the nine strategies. The ACMI scale was based on the following two-part definition of constructivism (p. 325):

1. *Constructivism is a belief that conceptual understanding in mathematics must be constructed by the learner. Teachers' conceptualizations cannot be given directly to students.*
2. *Teachers strive to maximize opportunities for students to construct concepts. Teachers give fewer explanations and expect less memorization and imitation. This suggests not only a perspective on how concepts are learned, but also a valuing of conceptual understanding.*

A 99% rater reliability was reached on both scales. The results show that teachers were able to adopt the constructivist perspective and practices. At the end of stage two only two teachers (4%) were assessed at level five on the ACMI scale. Of the 15 teachers that completed stage four, none reached level five at the end of level two, however, 11 (73%) reached level five by the end of stage four. On the strategies LoU scale, eight teachers (14%) were rated at level five after stage two. Only two of the 15 achieved level five after stage two, but 11 teachers (73%) reached the highest level after participating in levels three and four. Simon and Schifter (1991) concluded that teachers could learn new practices and views of learning consistent with the reform standards (NCTM, 1989), provided intensive in-service learning and support similar to ELM is provided. Providing teachers the opportunity to gain ownership of the curricula they teach was identified as the main professional development feature, which led to successful ELM outcomes. As teachers chose their own tasks, plan them, and reflect on student learning and engagement, they become more confident instructional leaders (Simon & Schifter, 1991).

2.4.2 Cognitively Guided Instruction (CGI)

The CGI program focused on using information gathered from research on students thinking to improve teacher knowledge, instructional practices, and student learning outcomes. The researchers' rationale was that research exists on how students apply a variety of strategies to solve addition and subtraction problems, but teachers do not make use of or have access to the research information (Carpenter, Fenema, Peterson, Chiang, & Loef, 1989). The CGI project employed a control (n = 20) versus experimental (n = 20) group methodology to identify differences. The treatment group participated in a four-week summer program while the control group was provided two separate two-hour workshops on problem solving. The CGI research team developed classroom-coding protocols used to observe each of the forty teachers classroom instruction. Toward the end of the school year after the summer professional development, teachers were asked to anticipate how their students would solve particular problems and this was measured against how students actually solved the questions on an end of the year student assessment. Additionally, teachers completed a survey gauged to capture their beliefs about student thinking and teaching of addition and subtraction. The data collected showed that CGI professional development, which focused on sharing students' thinking with teachers, affected their classroom practice and student learning outcomes.

The learning goal for the treatment group was to understand how elementary students employ various methods to solve addition and subtraction problems, and explore how the teachers could use this information to support their students in learning the same concepts. The researchers engaged teachers in analyzing student solutions. They sorted tasks into different groups according to possible student solution methods. After recognizing the features of the different types of adding and subtracting word problems, the teachers began to design their own instructional plans. The twenty teachers and CGI facilitator discussed instructional approaches, but none were prescribed. The four guiding instructional principles were (p. 505):

1. *Instruction should develop understanding by stressing relationships between skills and problem solving where problem solving is the organizing focus.*

2. *Instruction should be organized to facilitate students' active construction of their own knowledge with understanding.*

3. *Each student should be able to relate problems, concepts, or skills being learned to the knowledge that he or she already possessed.*

4. *It is necessary to continually assess not only whether a learner can solve a particular problem but also how the learner solves the problem.*

As designed, the teaching practices focused on what students were communicating and this implied that the teachers needed to think of ways to assess students' knowledge so that they could advance their understanding. In particular, teachers considered and planned questions to elicit children's thoughts. Finally, the experimental teachers examined curricula materials to learn if the various types of problems discussed were represented and to what extent. CGI conducted minimal follow-up during the instructional year meeting once in October and a project member was assigned to respond to all teacher questions.

The control group attended one two-hour workshop in September and a second one in February. These sessions focused on solving non-routine problems. Instructional frameworks were omitted. Teachers solved the problems and discussed the various solution methods. The group did discuss student thinking, but this focused only on how students might solve a particular problem not on sharing actual student work. The teachers also look at books that contained other non-routine problems and their own curricula for possible ways to encourage problem solving.

During classroom observations an elaborate coding systems was used to capture the teacher and target students' actions. One observer focused only on the teacher and other on the target students. The coders switched between observing and coding for 30-second intervals every minute. The observation protocol for the teacher included setting (student grouping: whole, small group, etc.), content, (i.e. number facts, word problems, etc.), expected strategy (i.e. recall, direct modeling, advanced counting, etc.), teacher behavior, process focus, and answer focus. The student protocol included setting, content, strategy used and lesson phase. These observations were used to measure the teachers' use of CGI instructional strategies. Some observation factors did not show a difference between the two groups. However, the CGI teachers administered significantly fewer memorizing tasks (low-level) and more problem solving tasks (high-level) than those in the control group. Secondly, the students in the CGI classrooms were more often presented problems to solve and the teacher more often listened to the processes students used than their counterparts. The final interesting finding from the observations was that CGI teachers spent less time reviewing concepts.

Teachers' knowledge of students thinking across three areas was assessed with interviews. The teachers were asked to predict target students' ability of number fact strategies, problem-solving strategies, and problem-solving abilities. For instance, in the first part of the interview the teacher was presented five number fact problems and asked to predict what strategy each of their target students would use. The teachers' responses were matched with how the students solved the problems. The CGI teachers outperformed their colleagues in all three areas, and were significantly better are predicting both students number fact and problem-solving strategies. There was not a significant difference with knowing how students would perform on complex addition and subtraction word problems.

The final CGI teacher instrument was four sets of 12 Likert style questions. The five-point Likert scale ranged from strongly agree to strongly disagree to determine a change in teacher beliefs between the control and experimental groups. The four sets of questions focused on the role of the learner, the relationship between skills, understanding and problem solving, sequencing of mathematics, and the role of the teacher, and was administered both pre and post treatment. On the first scale, the role of the learner a high score meant that the teacher believes the student needs to construct his or her own knowledge. A low score on the next section indicated that the teacher believes students first need to learn facts before they can engage in reasoning. The third set of 12 questions focused on how teachers should choose tasks where high score indicated that teachers believe they should be based on how students learn concepts. The final group questioned teachers belief about either engaging students in developing their own understanding versus presenting information to students. The CGI teachers changed their beliefs closer to the programs in all four categories between the pre and posttest. Both groups significantly improved with believing that students need to construct their own knowledge. While the control group reported a slight decrease in thinking students first need to practice basic skills before problem solving, the CGI significantly changed their belief that problem solving should drive student engagement in learning basic skills.

Carpenter et al. (1989) also report that these changes in CGI teachers' knowledge and beliefs also increased student-learning outcomes. The students, whom teachers participated in the treatment group, significantly outperformed their peers on basic skills and solving complex addition and subtraction problems. The researchers identify the fact that CGI teachers learned about the research on student thinking as a key ingredient to the programs success. As teachers gain a full understanding of research findings, this knowledge allows them to make more educated instructional decisions. Additionally, the research findings need to be practical and grounded in explicit students examples. Finally, Carpenter et al. recommend the mathematical content of professional development should be chosen judiciously to bridge student thinking with critical content.

A follow-up study (Fennema et al., 1996) to the original CGI project showed the sustainability of the program. Four years later, 18 of the 21 teachers improved their instructional practices. Initially teachers were identified as modeling routine procedures for students to reproduce and several years after the CGI, teachers engaged students in solving problems and conducting whole class discussions for students to communicate their thinking. Additionally, the change in instructional practices attributed to improved student outcomes. Students in classrooms where teachers changed their practice improved their students' ability to solve problems and these changes in instructional practices did not affect students' procedural knowledge. Therefore, the CGI project showed increased knowledge and assimilated reform instructional practices following the initial professional development sessions, and the teachers were able to sustain the CGI teaching principles four years later resulting in improved problem solving skills for the teachers' students.

2.4.3 Enhancing Secondary Mathematics Teacher Preparation (ESP)

The ESP project also identified the teacher and improving teachers' knowledge and practice as a way to increase student outcomes. The main program hypothesis is if teachers enact high-level tasks in pedagogically sound ways, then secondary students will improve their ability to reason mathematically. ESP was a professional development program to improve teacher's knowledge of cognitively demanding tasks and skill with implementing them.

ESP wanted to develop teacher leaders. The rationale was that the ESP teachers would mentor pre-service teachers and provide a classroom environment that brought to life the same practices the interns were learning about in their course work. So similar to the ELM project, ESP had stages of development over two years to educate the practicing teachers about enacting cognitively demanding tasks, and to define the teachers' role in supporting the teachers they were mentoring. In the first year, the teachers and the ESP facilitators met for six full days. At the end of the first year the group meet for a week to focus on their role as teacher leaders and mentors. In the second year, mentor teachers and the pre-service teachers assigned to their classrooms along with the ESP staff met for five half days. Boston and Smith (2009) studied the data collected from 18 mentor teachers as they participated in the ESP professional development during the first year.

The six full day professional development sessions were spread out during the school year where the teachers engaged in authentic teacher activities as recommended in the ELM (Simon & Schifter, 1991) and CGI (Carpenter et al. 1989) programs. The ESP teachers solved high-level cognitively demanding tasks, identified tasks based on their cognitive level, and analyzed instruction. In addition to working on the problems and activities posed in the professional development sessions, the teachers connected their thinking about solving and sorting tasks to their own curricula. For instance, they identified and planned activities to teach in their own classrooms. The teachers shared their classroom experiences during ESP sessions. Additionally, the ESP facilitators modeled the instruction they intended their teachers to utilize to create a collaborative learning environment, which supported the teachers in constructing their own knowledge. Finally, the ESP project also followed an explicit philosophy for their professional learning (p. 130):

1) *The importance of building professional development experiences on teachers' prior knowledge and beliefs*

2) *The assertion that change occurs as new conceptions of mathematics teaching and learning conflict with the teachers' prior knowledge and beliefs, and*

3) *The role of social interaction in stimulating and maintaining this type of conflict*

A total of 10 teachers were selected as a control group. This group did not participate in any of the ESP professional development, nor were they provided any other workshops. They were asked to participate only for research purposes to contrast their use and implementation of tasks.

Boston (2006) collected and analyzed data to identify changes in teacher's knowledge and instructional practices. During the first year of instruction data sets were collected at three different time periods (fall, winter, and spring) in addition to completing a pre-test, post-test, and post interview. A data set consisted of collecting instructional tasks for five consecutive days, collecting student work from three of the five tasks, and observing instruction in one of the five classes. All six of the professional development sessions were videotaped and course artifacts were collected. Finally, a post-test was administered and the ESP teachers were interviewed at the conclusion of the first year. The contrast group also completed the pre-test and each teacher in this group was observed one time at the same time as the spring data collection for the experimental group.

The pre and post-test was a task sort activity to identify teachers' ability to distinguish between high and low level tasks. The teachers scored the tasks as high, low or not sure and also provided a rationale. After scoring and providing rationales for all 16 tasks the teachers generalized their particular rationales. In other words, the teachers created their own general criteria that they could use for sorting any task. Boston (2006) used the TAG that was discussed earlier in this chapter to score responses. The results showed that there was a significant difference between the ESP teachers' post-test scores and the control group. In particular, improvement was shown in terms of teachers' ability to identify low-level tasks between the ESP pre and post-test.

Teachers collected all tasks that they engaged their students in solving for five consecutive days during the fall, winter, and spring seasons. The collected tasks included warm-up problems, main instructional activity, and homework assignments. Tests and quizzes were excluded. The Instructional Quality Assessment Academic Rigor (IQA AR-Math) rubric was used to score the instructional potential of each task, which is a five-point scale (0-4). Boston and Smith (2009) reported that teachers significantly improved their ability to choose high-level tasks over the course of the professional development. Additionally, it was noted that some teachers used a standards-based curriculum while others were in school that adopted a conventional text. The curricula showed no effect with choosing high-level tasks.

During the week of task collection, the teachers collected a full class set of student work from any three days. The teachers labeled the work as examples of low, medium, and high with respect to their expectation of quality solutions. Boston & Smith (2009) analyzed student-work among the three data collection periods to learn if the students engaged with the task at a high-level. The IQA AR-Math for potential was used to measure the level of potential of the mathematics task. The IQA AR-Math for implementation measured student engagement. Also a five-point scale (0-4), the implementation rubric is similar to the potential of the task rubric, but the implementation rubric addresses the actual student engagement as they solved the task. So two scores were given for each class set during each data collection the Fall, Winter, and Spring. Since three sets of student work were provided during a collection period, the potential and implementation scores were averaged separately. The two averages were compared to find out if the cognitive demands were maintained during instruction. Boston and Smith reported that students were afforded greater opportunities to learn at a high level between the Fall and Spring.

One classroom observation took place during the task collection week for each teacher. A total of 11 ESP teachers were observed three times during the school year and 10 control teachers were observed one. The lesson observer scripted various features of how the teacher enacted the class tasks including how the task was launched, the various interactions as the students worked on the problem, and the organization of the whole class discussion. The observations were scored on the tasks potential (IQA AR-Math), implementation (IQA AR-Math), and to score the factors of decline or maintenance the IQA Lesson Checklist was employed. The findings show that during implementation more tasks were maintained at a high level during the spring than the fall. The initial data collection of the experimental group showed a similar ability to select and enact tasks as the control group. However, a comparison between the control group and the ESP teachers' third implementation showed significant differences in both the potential of the tasks chosen and the actual instruction. Qualitatively, the IQA Lesson Checklist was useful in showing that the ESP teachers were seen holding students accountable for high-level outcomes, teachers questioning students thinking more, and pressing students to make more connections between the Fall and Winter data collections.

Boston and Smith (2009) provided evidence of teacher growth in terms of selecting and implementing high-level tasks based on the ESP professional development and research. The ESP is a learning program that chooses tasks closely aligned with teachers' practice and adopted the theory that change occurs when the learner experiences a cognitive conflict in a social setting. The research utilized the IQA in multiples ways to capture the teacher's growth. A unique feature of the ESP project was the use of the IQA on student work to study the implementation and selection of high-level cognitively demanding tasks.

ESP also studied the sustainability of their professional development program (Boston & Smith, 2011). The researchers found that two years after the professional development sessions and a full year of any professional develop support most teachers were still selecting high-level tasks and implementing them at a high-level. The researchers contributed the sustained affect to both the teacher's engagement as learners and mentors and the design of professional development project.

All three professional develop programs (ELM, CGI, and ESP) credited the success of increasing teacher knowledge and change in instructional practice on the professional development design features. These productive features are extracted and expanded upon in the context of the study herein in the next section to show how this study mimicked productive programs to also improve participants' knowledge for teaching mathematics.

2.5 SITUATING THE RESEARCH IN THIS CURRENT STUDY

This study draws on three areas of research: reasoning-and-proving, the role of mathematical tasks, and enacting and studying professional development that improves teacher knowledge and practices. The first two sections of this chapter (2.1 and 2.2) detailed what teachers need to learn to develop students ability to reason-and-prove, section 2.3 argued the importance of selecting high-level and section 2.4 explained how successful professional development programs were designed to improve teacher knowledge and practice. This last section of chapter 2 connects the features of the reasoning-and-proving course with the successful professional development programs to show why it too will be successful.

2.5.1 Features of Productive Professional Development Programs

This study draws on three areas of research: reasoning-and-proving, the role of mathematical tasks, and enacting and studying professional development that improves teacher knowledge and practices. The first two sections of this chapter (2.1 and 2.2) detailed what teachers need to learn to develop students ability to reason-and-prove, section 2.3 argued the importance of selecting high-level and section 2.4 explained how successful professional development programs were designed to improve teacher knowledge and practice. This last section of chapter 2 connects the

features of the reasoning-and-proving course with the successful professional development programs to show why it too will be successful.

A recent policy document reported on the results of eight professional development programs that impacted teacher instruction and or student outcomes (Wei, Darling-Hammond, & Adamson, 2010). All the studies occurred in either mathematics or science between 2004 and 2007. Wei et al. identified five design characteristics of the eight impactful professional programs as the following:

- *A strong focus on content and content-pedagogy in math or science;*
- *An annual duration ranging from 45 to 300 hours (or 9-12.5 graduate credit hours), and in most cases a design requiring more than 100 hours of engagement with both off- site (e.g., a two-week summer institute) and school-based components;*
- *Explicit links to, and thereby coherence with, the participants' school curriculum and organization;*
- *Elements of collective participation, bringing teachers together to engage in professional learning through coaching and mentoring by master teachers, lesson study with colleagues, additional training sessions focused on content pedagogy, and participation in learning activities with grade-level teams;*
- *Designs that are school-based and involve the schools as strong partners* (p. 6-7, 2010).

Comparing the five design characteristics with the three programs (ELM, CGI, and ESP) previously discussed shows some overlap and discrepancy. All three programs fell within the 45-300 hour time frame and focused on content and pedagogy. Furthermore, Carpenter et al. (1989) emphasized the first and third bullets. The addition and subtraction story problems were not only part of the curriculum, but existing research on student thinking on the subject was used to support teacher learning. In other words, focusing on the content is listed as impactful, but going a step further to include research on how students think about the content could be an expanded form of the first bullet. The ELM and ESP projects both included mentoring as additional steps to the professional learning. However, none of the three productive programs were school-based (last bullet). Finally, Wei et al.'s (2010) list is missing a connection to the greater research community in particular an organizing professional development framework and a learning theory to focus the instructional situations, which were explicitly described as instrumental in the three professional development programs.

2.5.1.1 Learning theory

Wei et al.'s (2010) five design characteristics includes the connection to content and content-pedagogy, but were not very specific about the types of teacher activities or a belief about how the teachers will expand their knowledge of content or pedagogy for teaching. The ELM program followed a social constructivist learning approach. The belief is that to conceptually understand mathematics the learner must construct the knowledge since an expert's conceptualization cannot be given directly to a novice (Simon & Schifter, 1991). As the teachers were engaged in activities that changed their perception of knowing mathematics a mental disequilibrium occurred which is labeled a cognitive conflict. The conflict is seen as a necessary part of the process in transforming teachers' view of mathematics in particular what it means to understand it. The mental conflict provides the learner with a reason to restructure their old thinking based of the new learning experiences. The ESP program also followed the constructivist theory to induce a cognitive conflict about the role tasks play in student learning along with the importance of not lowering the level of cognitive demand throughout the implementation of a task. The CGI researchers followed a related route to teacher change in knowledge and practice. The focus was on student thinking as they solved a variety of word problems. The teachers learned that listening to students as they explained their solution method and responding with appropriate questions is important in developing students understanding on mathematics. This realization promoted the CGI teachers to change their practice to accommodate their new understanding of how students think. In other words, the new understanding of how students make sense of situations conflicted with their previous belief what it meant to know addition. The teachers had to construct this understanding on their own to make instructional changes.

The reasoning-and-proving course also drew on these learning perspectives to expand teacher knowledge. The course aimed to promote disequilibrium in order to expand teachers understanding of how to evaluate and construct valid arguments. By engaging teachers in solving tasks and analyzing student solutions, they would reconstruct their view of what counts as proof. For example, a sequence of three mathematical tasks was specifically designed to create a cognitive conflict around the known misunderstanding that empirical examples count as proof (G. Stylianides & A. Stylianides, 2009). The typical trajectory is that learners write a generalization from a few examples and claim it as proof. The second problem in the three-task sequence encourages inductive reasoning, but after checking several cases the pattern fails. The learner reaches a conclusion that one must check more cases before generalizing. The final problem in the sequence forces the learner to reach a conclusion after checking many more examples only to learn that a counter example exists. The task sequence creates a conflict between what they previously believed proof to be and a new understanding that a generalization is not a proof and a formula cannot be trusted after only testing a few examples. Additionally, the learning occurs in a whole class setting so that the learners can discuss their thinking of the three-task sequence and, more generally, their understanding of what counts as proof.

Analyzing student arguments was included in the course design to also shape teachers' ability to identify proof from non-proof arguments. The design and intended implementation of the analyzing student solutions also contribute to a cognitive conflict and constructivist perspectives. The participants are asked to negotiate their personal criteria of proof as they label arguments. The student solutions chosen for teachers to evaluate were strategically selected to promote learners to rethink their view as to what is and what is not a proof. For instance, many students and teachers believe that proof needs to take on a specific form or be organized in a particular way. So solutions to proof tasks were chosen for teachers to critique that did not fit the conventional structure to reshape their mental image of what counts in some cases leading to a cognitive conflict. However, the reorganization was negotiated in a social context. Therefore, the reasoning-and-proving course drew on cognitive conflict and constructivist perspectives of learning.

2.5.1.2 Explicit use of organizing frameworks

The explicit course frameworks address learning what reasoning-and-proving means with respect to mathematical activities and planning reasoning-and-proving instructional tasks. The class was introduced to the Reasoning-and-Proving framework as they read the Stylianides (2010) article and labeled student arguments in a class activity. This framework highlights the various types of possible arguments students might produce, the range of activities that are involved with writing proofs, and provides a reminder for discourse with constantly questioning the learner about his or her conceptions of proof and why any presented argument may or may not count. On the same discourse thread, another purpose of the framework is that after recognizing and identifying an argument the teacher can choose questions to assess and advance students' thinking and move them toward deductive reasoning. Lannin (2005) found that promoting generic arguments is a productive path away from empirical, which is listed in the framework.

A second organizing framework explicitly discussed in the course is to promote the planning of reasoning-and-proving. Instead of presenting this framework, the teachers read various articles on each of the three constructs: Task, Tools, Talk. From the articles the teachers list the essential characteristics of each "T", and the class creates an encompassing shared understanding. For instance, the task is expected to problematic and leaves students with a type of learning residue. The residue could either lead to students learning a new mathematics concept or a mathematical process. Tools can be diagrams, algebraic symbols, or any other instrument useful to help students access the mathematics. For example, when the teachers engaged in an activity that required circles the facilitator provided them with a sheet of paper with the circles already drawn. Since drawing circles was not the goal of the lesson, the tool was useful with assisting the learners with focusing on the mathematics. In other words, the course facilitator also modeled the framework. The teachers applied their understanding of this planning framework as they engaged in various activities before using it to plan a complete lesson.

The design and implementation of the reasoning-and-proving course included most of the bulleted list and all three of the features identified in the three productive professional development programs: focus on content, explicit use of frameworks, and implementation of authentic activities of practice. The only feature missing from the reasoning-and-proving course design was a strong the connection to a school. None of the pre-service teachers in the course were placed in the same school. The course aimed to build teachers' capacity to implement high-level reasoning-and-proving tasks, which includes learning to reason-and-prove, select tasks, and understand pedagogical practices to support student development. To increase teachers' knowledge, teacher educators need to engage teachers in authentic activities of practice (Boston & Smith, 2009; Putnam & Borko, 2000; Simon & Schifter, 1991). The student in the reasoning-and-proving course solved high-level math tasks, discussed episodes of practice, and planned instructional lessons. As discussed in section 2.1 of this chapter, student understandings were utilized in the design of the activities. The teacher tasks will be explained further in chapter 3.

2.5.2 Grounding study in a theoretical model

Hanna (1995) identifies the main challenge to increasing students experiences with proving is the lack of opportunities across K-12 content. Phase one of the Mathematical Tasks Framework (MTF) includes selecting or designing high-level tasks. Given the limited number of tasks found in high school textbooks outside of geometry (Johnson, Thompson, & Senk, 2010), secondary teachers will need to do more *designing* of tasks in order to provide student opportunities to engage in reasoning-and-proving across all secondary curricula. From experience during the first iteration of the reasoning-and-proving course materials (summer 2010) it became clear that teachers struggled to select and or modify tasks. Since a mathematical task sets the stage for the work of teaching and student learning (Doyle, 1988; Hiebert et al., 1996; Smith & Stein, 2011; Stein et al., 2010), teachers need to improve their skill selecting and or modifying reasoning-and-proving tasks (unpacking phase one of MTF as shown in figure 2.8). Furthermore, learning to modify reasoning-and-proving tasks cannot be taught to teachers that do not know what is proof or are unable to produce a valid argument.

Just as it is believed that different types of mathematics knowledge for teaching is needed to move between selection and set up of a mathematics task (Stein et al., 1996), similar knowledge is needed to modify a task to include reasoning-and-proving (as shown in the circle in figure 2.8). Narrowing the scope to reasoning-and-proving tasks, the hypothesis is that the R&P goal the teacher chooses, the teachers' knowledge of R&P, and knowledge of students about R&P all impact the modification process. Identifying a mathematical goal should not be overlooked (Smith & Stein, 2011), since the solutions student write provides the teacher with information about whether the goal was or was not accomplished (Hiebert, Morris, Berk, & Jansen, 2007). Identifying a learning goal was developed throughout the course as the participants read narrative cases and were directly asked what they believe students learned. Secondly, teachers' ability construct arguments may influence their skill with selecting or modifying tasks. Throughout the course, including the interviews, the teachers were asked to write eight proofs and then think about why the argument is or is not a proof. Finally, teachers need to know what typical students do when asked to write a proof. Knowledge of student solution methods and prior knowledge in the domain is useful in selecting appropriate tasks (Carpenter et al., 1989). While the course aimed to build teachers ability to identify a goal, write and evaluate proofs, understand student thinking, the participants were explicitly taught how to modify tasks.

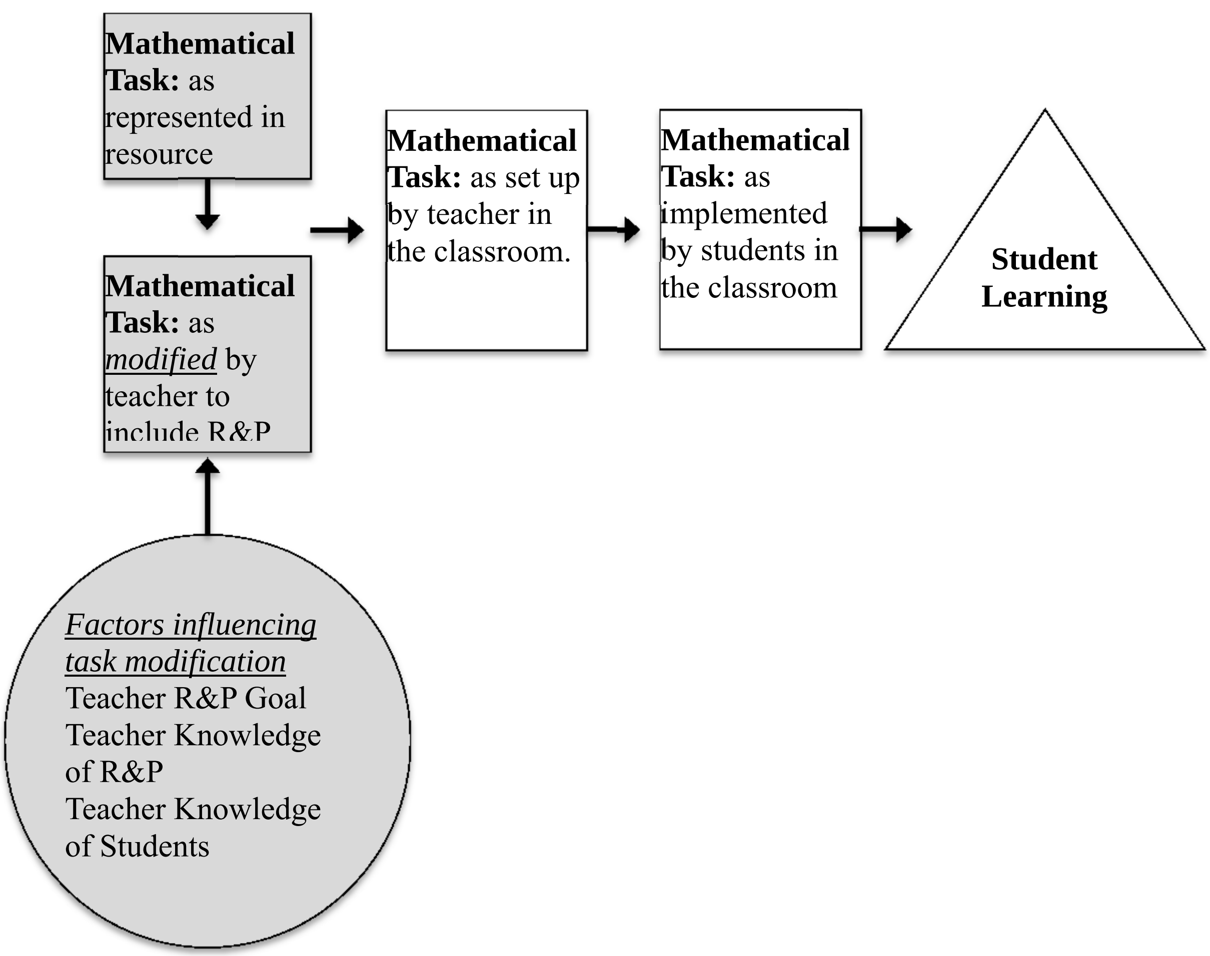


Figure 2.7. Unpacking Phase 1 of MTF and identifying factors that influence task modification

Deciding on a mathematical goal along with an increased knowledge of reasoning-and-proving including student thinking are factors that may contribute to task modification. Explicit instruction followed a '*to with by'* model which is a form of what Collins, Brown, and Newman (1989) call a "cognitive apprenticeship" or scaffolding. The three parts include a modeling (*to*), coaching (*with*), and then a fading (*by*) of support to promote development. The participants were shown typical tasks along with a modified version of the same task. In other words the modifications were shown *to* the participant to think about how each task was altered to include reasoning-and-proving. After a series of such activities, the class derived a set of modification principles. The principles were then applied to a new set of unmodified tasks and the facilitator modified the tasks *with* the participants. Finally, the participants were asked to modify tasks from their curricula *by* themselves. During each of the three phases the concept of solving the tasks as both a knower of mathematics and a learner to focus on the student perspective. Therefore, knowledge of R&P and student thinking were developed along with explicit learning with how to modify tasks. Again, this was a main focus of the course since high school curriculum provides limited opportunities for students to reason and prove. Supporting teachers with selecting and or modifying their current curricula is intended to increase the number of reasoning-and-proving experiences for students, which currently a challenge (Hanna, 1995).

3.0 CHAPTER 3: METHODOLOGY

This design-based research study investigated teachers' developing understanding of reasoning-and-proving during their participation in a methods course focused on reasoning-and-proving in secondary mathematics. Teachers' developing understanding was evaluated through the examination and analysis of structured interviews, work produced during the course, and artifacts provided by teachers from their classrooms in the academic year following their completion of the course. The specific research questions that are the focus of analysis in this study are:

1. *How do pre-service teachers' conceptions (i.e. purpose of proof, what counts, proof in secondary courses) of proof change over the duration of a course focused on reasoning-and-proving?*

2. *To what extent do pre-service teachers construct valid and convincing arguments when prompted to write proofs over the duration of a course focused on reasoning-and-proving?*

3. *To what extent do pre-service teachers improve their ability to distinguish between proof and non-proof arguments created by students over the duration of a course focused on reasoning-and-proving?*

4. *To what extent do pre-service teachers improve their ability to select and or modify reasoning-and-proving tasks for students over the duration of a course focused on reasoning-and-proving and during their first year in the classroom?*

The following sections describe the context of the intervention including the participants, the collected data, and how the data was coded and analyzed. The first section explains the course that was implemented to increase the participants understanding of reasoning-and-proving, and how this course fits into other courses that are part of the teacher-credentialing program the participants completed. The second section in this chapter details the data collected from the course and interviews. The third and final section explains how the collected data was coded and analyzed related to each of the four research questions.

3.1 CONTEXT OF R&P COURSE

This study focused on the extent to which learning occurred in a course intended to develop teachers' knowledge related to reasoning-and-proving. The course engaged teachers in writing proofs, critiquing student work, analyzing narrative cases, selecting and modifying tasks, discussing mathematical and pedagogical issues, and reflecting on their own learning. A total of 10 students enrolled in the course, of which nine participated in the study. The following two sections explain the course and those who participated in it.

3.1.1 The reasoning-and-proving course

The reasoning-and-proving course included 12 (3 hour and 15 minute) sessions equally sequenced over a six-week time period starting on May 10, 2011 and concluding on June 16, 2011. The course was designed around a set of materials developed under the auspices of NSF-funded **C**ases **O**f **R**easoning and **P**roving in Secondary Mathematics (CORP) project. The purpose of the CORP project is to design curriculum materials that can be used in the professional education of pre-service and in-service secondary mathematics teachers. Three key questions guided the development of the materials and the course:

1. *What is reasoning-and-proving?*
2. *How do high school students benefit from engaging in reasoning-and-proving?*
3. *How can teachers support the development of students' capacity to reason-and-prove?*

The guiding questions were embedded throughout the course activities. The first and second questions were directly asked of the teachers, and the third question was embedded in the course frameworks. The implementation of the activities included individual work, pair-share, small group and whole class discussions. The next two sections will detail the course activities and how they were enacted.

3.1.1.1 Course activities

The course map (as shown in Figure 3.1) outlines the six key ideas explored, and the types of activities enacted across the class sessions. The numbers across the top of the map signifies each of the 12 course meetings. The figures in each column reflect the nature and sequence of the activities enacted during each class period. The six key ideas explored, listed below the map are, for the most part, grouped as consecutive activities and tagged with a symbol. For instance, in the map the third and fourth activity on day one and the first activity on day two have check marks. All three of these activities explore the same key idea of *Motivating the Need for Proof* as indicated in the key below the map.

The shapes identify the type of teacher learning activity. For instance, the rectangles identify the instances when the participants were asked to solve a mathematical task. The exploration of five of the six key ideas includes a mathematics task (rectangle). The activities such as analyzing student work (hexagon) or a narrative case (oval) relate to the mathematical task that participants had previously solved related to the key idea. The only exception is that the third idea does not start with a proof task. The narrative case of Nancy Edwards is the "Odd +

Odd = Even" task solved in the second unit. So teachers explored a variety of practice-based activities in order to develop their understanding of a key idea and in most units a task was solved first.

Reflection on learning (cloud) was an ongoing process. The teachers also read articles about the course frameworks to develop a shared understanding for selecting and planning reasoning-and-proving tasks. The reasoning-and-proving (R&P) framework, discussed in unit two, was introduced when the teachers read Stylianides (2010) at the beginning of class three. Additional activities were used to develop their understanding of each of the terms in the framework. Throughout the course the R&P framework was utilized to identify the potential of tasks and to sort arguments. The homework assignments are listed below the horizontal grey strip across the bottom of each class period.

The arrows represent connecting to practice (CtoP) activities that were intended to support the participants in applying course concepts to actual teaching practice, namely the planning of a reasoning-and-proof lesson. CtoP are not specifically related to one key idea, which is why they are not marked with a symbol. They were inserted throughout the course to provide the participants an opportunity to apply the course concepts, and serve as a formative assessment for the instructor.

Finally, there are six other figures in the map without an identified connection to a key idea (two on day one, one on day four, and three on the last class meeting). These activities were intended to gather information based on the participants' current understanding or thinking about previous activities or to share information about reasoning-and-proving in general. The meaning of the shaded activity shapes will be explained in the data collection section 3.2.2.

The six key ideas aim to reach larger goals and the individual activities within each key idea promote smaller goals to foster teachers' development with implementing reasoning-and-proving tasks in their classrooms. For instance, the first key idea is *Motivating the Need for Proof,* which focuses on the shortcomings of empirical arguments or the use of examples to generate a generalization. The sequence of three tasks (Squares, Circle & Spots, and Monstrous Counterexample) (G. Stylianides & A. Stylianides, 2009) presses on the participants' mathematical knowledge of the limitations of developing a generalization from any set of examples. The main question is: how many cases must one check to determine the truth of a generalization? The realization is that no number of examples is enough, which is why proof is needed. Within the same unit or key idea, two narrative cases describe how two different teachers implemented the same set of three tasks with their students. One of the narrative cases (Kathy) provided the participants with an exemplar case with how to support students with learning the limitations of an empirical argument. The other teacher (Charlie) inhibited his students in learning why empirical arguments are not proof by taking over the thinking for them. In other words, the overall goal of the unit is to *Motivate the Need for Proof* and the first activity (sequence of three tasks) is intended to support the participants' mathematical knowledge of the limitations of examples and the second activity (analyzing narrative cases) focuses on building the participants pedagogical knowledge with how to implement such tasks so that students are supported in reaching the same mathematical goal that proof is needed. The other five units are similar in that the main idea is general and the individual activities aim to build mathematical or pedagogical understanding related to proofs.

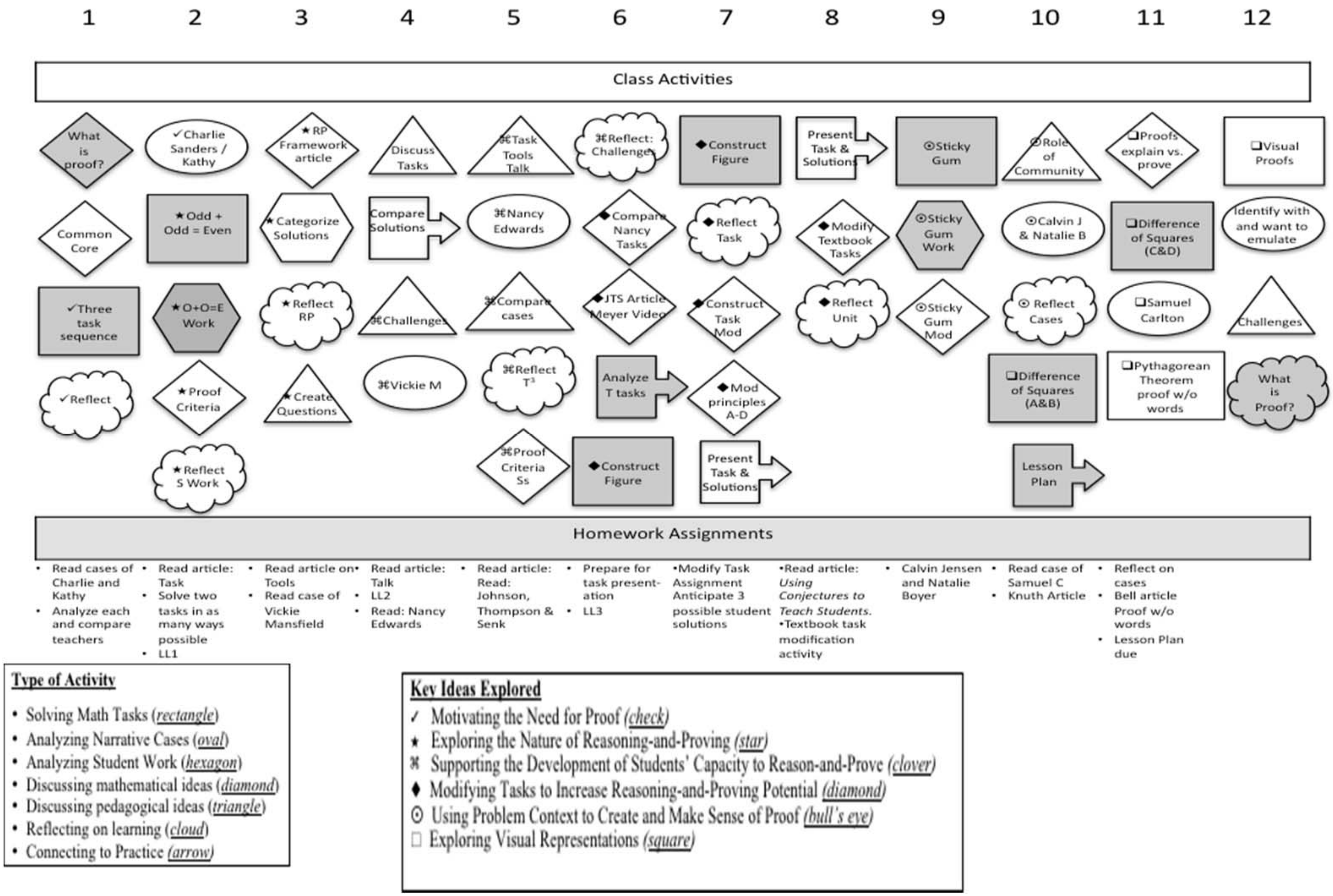


Figure 3.1. Course map summer 2011

3.1.1.2 Facilitation of activities in reasoning-and-proving course

The author of this study was the instructor for the reasoning-and-proving course. The implementation intended to engage the students in constructing individual and a collective understanding of the key ideas related to reasoning-and-proving as described in the course map. For instance, the participants worked on each mathematics task individually before sharing each other's thinking. Then pairs of students would share and connect their thinking with the class community. Most discussions were held at the whole class level, while the reflections were individual recordings. The reflections were not typically shared instead the instructor collected them to make future instructional decisions. One goal of implementing the course activities was to model pedagogical practices intended for the participants to enact in their own classrooms (Simon & Schifter, 1991).

3.1.2 The Participants

The students in the course were the participants in this research study. A total of 10 students enrolled in the Master's level course at the University of Pittsburgh. There exist two types of Masters programs for mathematics education at the university. The Masters of Arts in Teaching (MAT) is a credentialing yearlong program for those with an undergraduate degree in mathematics (or equivalent). The cohort group in the MAT program spends their days at their assigned secondary school with a mentor teacher and evenings taking courses at the university. The Masters in Education (MEd) is designed for teachers that have already earned a teaching credential, but are seeking an advanced degree in mathematics education for various reasons. Eight of the ten students in the class were part of an MAT cohort. This cohort group was considered unusual based on their strong mathematics aptitude and overall ability to discuss and comprehend new pedagogical concepts. One student, not part of the MAT cohort, earned a secondary credential and MEd over two years. All nine of these students were enrolled in their final course prior to earning their respective degrees. The tenth student did not participate in the study. Nine of the participants are female and one male. The nine participants completed the class activities and their notebooks were collected at the end of course and photocopied.

Since this was the participants' last course in the program they had many opportunities in previous courses to engage in similar work. For instance, the participants were use to solving open-ended tasks and discussing their solutions. The participants also learned the difference between low and high level tasks. In particular, the participants had previously solved pattern tasks and shared the different ways one could generalize the pattern, analyzed episodes of teaching that highlight the implementation of pattern tasks, and analyzed students' thinking related to these tasks.. It is likely that their experiences in previous courses prepared these participants for the curriculum expectations and instructional style they encountered in the reasoning and proving course.

In addition to the courses the participants completed during their credentialing program, those in the MAT cohort spent the school year teaching and observing a mentor teacher in a secondary classroom. This afforded the MAT students the opportunity to experiment as a teacher with ideas they were learning at the university. For instance, two participants engaged a group of students in solving a task they selected for the reasoning-and-proving course to gain a better insight into how typical students may engage with the problem. The MEd student was in a classroom for one semester and she spoke about her teaching experience as being limited. Therefore, since most of these participants spent a year with a mentor teacher and students, they were provided an advantage with engaging students in the practices they learned as university students.

Prior to the start of the reasoning-and-proving course the registered students were contacted for interview purposes. All nine participants were interviewed three times: 1) prior to the start of the course, 2) between the fourth and fifth classes, and 3) after the last class meeting. (The 10th student registered late for the course, which is why she did not participate in the study.) All the interviews were audio recorded, which were then transcribed. The participants were compensated $24 per interview.

At the conclusion of the course all participants who were employed as secondary teachers mathematics teachers were contacted. Of the seven who secured teaching positions, six were contacted and all six agreed to participate in the follow-up study which involved collecting artifact packets around R&P lessons they implement during their first year as teachers for which they would be compensated $100 for each returned packet. The seventh participant moved out of the area and new contact information was not available. Five of the six that agreed to the follow-up study where part of the MAT cohort, all six earned an undergraduate degree in mathematics, and all are female. In the end, however, only two teachers (1 MAT, 1 MEd) actually collected materials. It is not clear why the other four teachers did not choose to submit artifact packets. To promote clarity throughout the rest of this document, "participants" will be used solely to refer to the subjects in this study as opposed to pre-service teachers, teachers, or students. Additionally, all nine participants are named using pseudonyms.

3.2 DATA COLLECTION

Three main data sources were used to answer the research questions: interviews, course notebooks, and task packets completed by participants who took part in the follow-up study. The research questions will be answered using the data as shown in table 3.1.

Table 3.1 Collected data for analysis

Research Questions	Data Sources	Analyzed Data
1. How do pre-service teachers' conceptions (i.e. purpose of proof, what counts, proof in secondary courses) of proof change over the duration of a course focused on reasoning-and-proving?	a. Interviews b. Notebooks	a. Open-ended questions in each interview b. Opening activity; final reflection
2. To what extent do pre-service teachers construct valid and convincing arguments when prompted to write proofs over the duration of a course focused on reasoning-and-proving?	a. Interviews b. Notebooks	a. Solutions to three R&P interview tasks b. Solutions to five course tasks
3. To what extent do pre-service teachers improve their ability to distinguish between proof and non-proof arguments created by students over the duration of a course focused on reasoning-and-proving?	a. Interviews b. Notebooks	a. 14 student arguments b. 18 student arguments from two problem sets
4. To what extent do pre-service teachers improve their ability to select and or modify reasoning-and-proving tasks for students over the duration of a course focused on reasoning-and-proving and during their first year in the classroom?	a. Interviews b. Notebooks c. Task packets	a. First and third interviews (5 participants brought a task to interview 3) b. Two tasks selected during the course c. Two teachers enacted five tasks each during their first year as teachers (2011-2012)

The timing of the data collection is shown in figure 3.2. The twelve solid vertical lines represent the class meetings distributed across the six weeks. The dotted lines indicate when the interviews took place. Notebooks include participant responses to course activities from 12 class meetings as indicated below the timeline. The artifact packets were collected between November 4th and March 1, 2012. The following three sections will further explain the interviews, task packets, and notebooks.

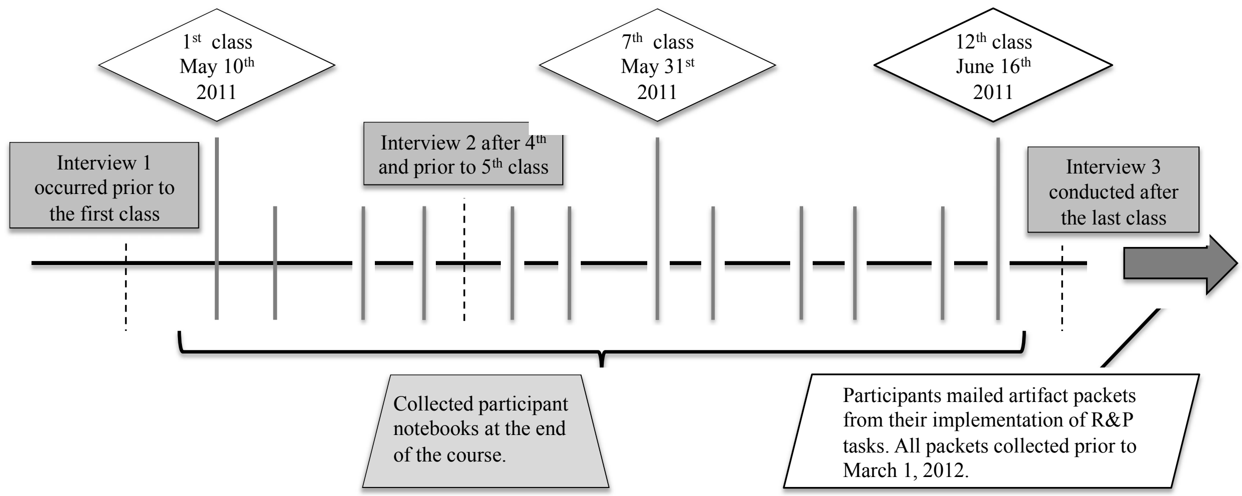


Figure 3.2. Data collection timeline

3.2.1 Structured interviews

As shown in figure 3.2 the participants were interviewed prior to the first class, between the fourth and fifth classes, and after the conclusion of the course. The participants were contacted and appointments were arranged for the interviews. At the time the interviews were scheduled, only nine students were enrolled in the course. All nine students were interviewed three times. The multiple part interview design, implementation of interviews, and questions in the protocols were based on previous proof interview studies (Chazan, 1993; Knuth, 2002a, 2002b; Morris, 2002; Smith, 2006; Solomon, 2006). Three members of the CORP project team conducted and audio recorded the 27 student interviews. The researcher / instructor did not interview students and not revealed until the course was over.

The author of this study and the three interviewers collaborated in the development of the three interview protocols (Appendix A). The three interview protocols included a consistent design of at least three parts. In the first part, participants responded to open-ended questions, followed by solving a reasoning-and-proving task, and then by analyzing student work to the task they solved in part two. Interviews one and three each had a similar fourth part, which included the participant sharing a reasoning-and-proving task they selected prior to the interview.

Open-ended Questions: Part 1 Interview. The first part of each interview engaged the participant in explaining their understanding of reasoning-and-proving through open-ended questions. The interview questions were designed to probe the participants evolving conceptions of proof and the impact the course had on their thinking. Since the first interview was conducted before the course started, the first question (see column one of table 3.2) was intended to gather information about the participant's previous experiences. The next three questions focused on learning the teachers' conceptions of proof in the discipline of mathematics. The final two questions in interview one press the participants to think about proof in secondary schools. The second and third interviews also include questions about the participants' conceptions and ask about the impact the course was having on their thinking. The participant responses to these questions were analyzed to answer the first research question.

To promote reliability and validity among the three interviewers and across the interviews, the interviewers were asked to only pose the stated questions and in the order presented in table 3.2. The interviewers did ask generic follow-up questions to encourage a participant to talk more about a particular topic or to clarify responses. For instance, an interviewer asked questions such as: Can you say more about that? I am not sure I understand, can you explain that for me again?

Table 3.2. Opened-ended interview questions

Open-ended questions (Part 1)		
Interview 1	**Interview 2**	**Interview 3**
What experiences have you had with proofs –as a student in high school and college and as a mathematics teacher? What does it mean to prove a statement? What should be included in a proof? What should or could a proof look like? What role do you think proof should play in the secondary mathematics classroom? Which courses in the secondary curriculum should or could include work on proofs?	What do you think is required for an argument to count as proof? Why? How, if at all, has your understanding of reasoning and proving changed over the last four classes? What specific activities do you believe have most helped YOU to better understand reasoning-and-proof? What, if anything, about reasoning-and-proof still is unclear or confusing? How has the course influenced your thinking about teaching reasoning and proving in your classroom?	1a. How, if at all, has your understanding of reasoning-and-proving changed over the past six weeks (12 classes)? That is, what is it you understand now that you did not understand prior to taking this class? 1b. What specific activities do you believe have most helped YOU in better understanding reasoning-and-proof? 1c. What, if anything, about reasoning-and-proof still is unclear or confusing? 2a. How has the course influenced your thinking about teaching reasoning-and-proving in your classroom? 2b. What specific activities do you believe have influenced YOUR thinking about teaching students to reason-and-prove? 2c. What, if anything, about teaching reasoning-and-proof still is unclear or confusing?

Solve R&P Task: Part 2 of Interview. Part two of all three interviews engaged the participants in solving a reasoning-and-proving task (as shown in Table 3.3). The problems were chosen based on several factors such as: the opportunity for multiple solution paths, accessible content relating to the secondary school mathematics content, and the task either explicitly called for a proof or requested a convincing argument. The first problem (1st column table 3.3) was adapted from an interview study Morris (2002) conducted with undergraduate students. It is a typical number theory problem, which allows for the generation of examples or to draw a diagram using a square and rectangle without explicitly calling for either. The task used in the second interview (2nd column in table 3.3) was adapted from the Interactive Mathematics Program (IMP) high school curriculum. It explicitly promotes the use of a diagram, and the generation of a conjecture. Being a pattern task, it is more accessible for the participant to produce a generic argument proof (Lannin, 2005). The contextual task (3rd column in table 3.3) used in the third interview was adapted from a middle school standards project (Achieve, 2002). The task promotes multiple mathematical representations and an opportunity to generate examples. While the task does not use the word prove or proof, it does ask for a justified argument. This is the only problem selected where the proof is a counterexample.

After the participants were provided sufficient time to solve a task, the interviewers asked two follow-up questions addressing the validity of their argument and understanding of generality (as shown at the bottom of each column in table 3.3). The first question was to find out if the participant believed that he or she wrote a proof. The purpose of this question was to gain access into their thinking about their established criteria of proof. For instance, Chazan (1993) and others reported that just because a student constructs an argument to a proof task does not mean he or she believes their solution is a proof. The second questions differed across the three interviews such as: generality of proof (interview 1), multiple proof methods (interview 2), and an opportunity to verbalize a proof if they did not believe they wrote one (interview 3).

The participants were provided a task sheet, which included the task only and not the follow-up questions. While the participants were solving the R&P task, the interviewers asked the participants to talk through their thinking to understand how they were approaching the problem. The interviewers did not suggest solution paths, nor did they try to advance a participant's thinking toward a proof. Instead the interviewers only asked clarifying questions to capture progress as each participant worked on a solution. The two follow-up questions were asked in the order listed in table 3.3. The participant responses to the reasoning-and-proving tasks were analyzed to answer the second research question.

Table 3.3. Reasoning-and-proving tasks: Part 2 of each interview

Interview 1 R&P Task	Interview 2 R&P Task	Interview 3 R&P Task

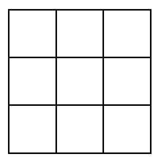
<table>
<tr>
<td>1.) Prove that for every counting number n (1, 2, 3, 4 …), the expression $n^2 + n$ will always be even.

Provide time for interviewee to prove the task. Then ask:

2.) What about your solution makes it a proof?

3.) Do you think that there is a counting number n which would cause the expression $n^2 + n$ NOT to be even? Why or why not?</td>
<td>The diagram below shows the frame for a window that is 3 feet by 3 feet. The window is made of wood strips that separate the glass panes. Each glass pane is a square that is 1 foot wide and 1 foot tall. Upon counting, you will notice that it takes 24 feet of wood strip to build a frame for a window 3 feet by 3 feet.

Determine the total length of wood strip for any size square window. Prove that your generalization works for any size square window.

3ft-by-3ft


Provide time for the interviewee to create a proof. Then ask:
What about your solution makes it a proof?
Can you think of other possible ways to prove that your generalization works (without writing it out)?</td>
<td>Long-distance Company A charges a base rate of $5 per month, plus 4 cents per minute that you are on the phone. Long-distance Company B charges a base rate of only $2 per month, but they charge you 10 cents per minute used.

Keith uses Company A and Rachel uses Company B. Last month, Keith and Rachel were discussing their phone bills and realized that their bills were for the same amount for the same number of minutes. Keith argued that there must be a mistake in one of the bills because they could never be the same. Rachel said that the phone bills could be the same.
Who do you think is right, Keith or Rachel? Why?
For any two phone plans, is there always a number of minutes that will yield the same cost for both plans? Provide an explanation to justify your position.

Provide time for the interviewee to create a proof. Then ask:
Is your solution a proof? Why or why not?
If not, what would it take to make it a proof?</td>
</tr>
</table>

Analyze student work: Part 3 of interview. The participants analyzed constructed solutions based on the task they solved in part two in the third part of each interview. The solutions were carefully selected to include all the argument types in the R&P framework. For instance, every solution set included an empirical argument. Additionally, the types of proofs included both generic arguments and demonstrations, and were designed to include a variety of representations such as the use of diagrams and narrative language. All 14 constructed arguments across the three interviews are listed in appendix A, and table 3.4 categorizes the types of argument with descriptive language to highlight the variety of the solutions the participants were asked to analyze. The student solutions to the task in the first interview were modified from Morris (2002) where she interviewed pre-service elementary and middle school teachers. Two solutions are empirical arguments, two are proofs, and one is a rationale (non-proof) since all statements are not clearly developed. The second and third interviews include a variety of types of proofs and non-proof arguments.

During each interview, the interviewer presented the participant with the collection of arguments and asked two or three questions. The first question was: which of the solutions are and which are not proofs and why? The participant would review each argument individually and then provide a rationale for their classification. Then after providing their rationale the interviewer asked which argument was most convincing. Finally, interview two included a third question about supporting a student in order to improve their argument. The questions that accompanied the work samples in interview one and three specifically focused on the validity of the five student solutions. While the participants were analyzing sample solutions during each of the three interviews, the interviewer encouraged them to write down comments or to talk through their thinking. The participant responses to the student work were analyzed to answer research question three.

Table 3.4. Argument types for each student solution: Part 3 of each interview

Solution	**Argument type with explanation**		
	Interview 1	**Interview 2**	**Interview 3**
A	Proof (demonstration): elegant novel response using narrative language and algebraic symbols	Empirical: constructs a generalization from a few examples without justifying why the pattern will always be quadratic.	Not a valid argument. An incorrect narrative response.

B	Empirical: Uses small and large numbers	Proof (generic example) uses a particular case to generalize to any size window.	Proof: counterexample. Provides a narrative general and a specific counterexample.
C	Empirical: A single example to generalize even numbers and a second single example to generalize odd numbers.	Proof (generic argument) uses a particular case and generalizes it different from solution B.	Not valid. Provides a convincing response without attending to the question.
D	Proof/rationale: Could be considered a rationale since the argument makes assumptions such as: If n is even, then n2 is even.	Empirical: finds a pattern by extending the diagram	Not valid. Again not responding to the question.
E	Proof (generic argument): relies on diagram (specific cases) to generalize for all cases	N/A	Proof: provides a general counterexample argument.

Select a reasoning-and-proving task: Interview Part 4 of interview one and three. The first and third interviews included a fourth part which was for the participant to select a reasoning-and-proving task prior to the interview that they believed provided students an opportunity to reason with or prove a mathematical statement. The interviewer asked why the participant selected the particular task. The task the participant brought to the first interview was revisited at the final interview along with the new task the participants were asked to bring. The participant selected reasoning-and-proving tasks were analyzed to answer research question four.

Each interviewer was assigned three participants for all three interviews. The interviews ranged in time between 30-60 minutes. All 27 interviews were audio recorded and transcribed and all written work was collected. The transcripts and written work were analyzed in order to address all four research questions as shown in table 3.1.

3.2.2 Course Notebooks

The CORP project materials, which were used in the course, include both facilitator resources and teacher handouts. The handouts varied depending on the specific activities as outlined in the course map (figure 3.1). On the first day of class the participants were given a binder to store all of their notes and materials. This included the handouts related to each of the key ideas, and the narrative cases, frameworks, and articles. The binder was equipped with 12 hard stock sheets labeled day one through day twelve to separate their work by class meeting. All of the course handouts were hole punched so that the students could file their work in their notebook binders at the conclusion of each class period. On the last day of class the binders were collected, photocopied, and subsequently, returned.

The shaded activities shapes in figure 3.1 are the activities that were analyzed to answer the research questions. The participant responses to the first and last activities both labeled "what is proof?' were analyzed to answer the first research question. The open-ended question was asked as the first and last course activities to gauge the breadth at which the participants could communicate their understanding of proof.

The participant solutions to the shaded rectangles (solving tasks) contribute to answering the second research question. These activities provided the participants an opportunity to improve their reasoning-and-proving skills with secondary content. Furthermore, the reasoning-and-proving course tasks were specifically chosen to expose the participants to a variety of mathematical situations, which allow for multiple solution paths. The rationale was that if teachers were capable of solving problems using a variety of representations and solution paths, then the teacher would hold a positive disposition toward implementing reasoning-and-proving tasks and be more successful with interacting with students' multitude of approaches.

The participants completed handouts in which they analyzed student work (two shaded hexagons), which will contribute toward answering research question three. These analysis activities provided the participants with examples of how students may solve problems, strengthen their criteria for proof, and begin to think about how to support students thinking. These are instructional skill teachers need to develop in order to implement high-level tasks successfully (Smith & Stein, 2011).

The two shaded arrows represent the two tasks the participants selected for planning purposes and will be coded and analyzed to answer question four. To reach the goal of integrating proof throughout the secondary curricula, teachers need have a broad conception of proof, develop an ability to construct proofs in multiple ways, understand how to support student thinking, and build their skill with selecting and or modifying tasks to met their instructional goals. The teacher handouts for the five reasoning-and-proving tasks and student work analysis handouts including the student work are attached (Appendix B).

3.2.3 Artifact task packets

During the 2011-2012 school year, six of the seven participants who secured a secondary teaching position volunteered to collect classroom data around their instruction related to reasoning-and-proving. Participants were asked to complete an artifact task packet for each task, which includes selecting a reasoning-and-proving task and a modified version if applicable, nine pieces of students' work organized into three categories (below expectations, met expectations, exceeded expectations), and any materials created in preparation for the lesson or during implementation. Additionally, the participants were asked to complete a task cover sheet and a background sheet.

The material found in appendix C details the information that was emailed to each of the six participants on November 4th. The first page of the document lists the requested data for each lesson packet they return. The second page is the task cover sheet, which will be completed for each task. The teachers were asked to collect student work on any reasoning-and-proving tasks they implemented or will implement prior to March 1, 2012. If they modified any task from their text or an outside resource they were asked to send both the original and modified versions of that task. The participants were also asked to provide photo copies of three pieces of student work that *exceeded their expectation* (1: EE), three pieces that *met their expectation (2: ME)* and three that they regarded as *failed their expectation (3: FE)* for a total of nine pieces of student work in each task packet. In the task cover sheet that the participants completed for each task they explained what their expectation was for each problem such as producing a proof, non-proof argument, make a generalization, etc. Furthermore, the teachers were asked to send any and all documents that they prepared to support their preparation for implementation of the task. Finally, a background sheet was included to be complete once to explain how they perceive the support they are provided by their colleagues and administration related to enacting reasoning-and-proving activities. Each participant was mailed an initial five envelopes affixed with postage and mailing addresses. Teachers had the option to send more than five task packets if appropriate. However, in the end only two of the six participants submitted task packets and only the enacted tasks and task cover sheets were used in the analysis of this study.

The rationale for the task packet is to gain a greater understanding of the selected and implemented R&P tasks opposed to just asking for the activity sheets. The type of tasks teachers select and enact affords students particular opportunities (Doyle, 1983). The cover sheet allowed for the participant to both foreshadow anticipated outcomes and reflect on the class engagement.

The student work supplies credence to the task being implemented and provides the participant an opportunity to compare their anticipated outcomes with the actual student solutions. The classification of student work provides evidence for student expectations. Finally, the background sheet was designed as to not make false claims about a participant based on how often they select R&P tasks. Since all of these teachers were new to their schools, it would make sense that they adhere to school and or district norms. The background sheet is intended to give the participant an opportunity to explain the extent to which they felt supported by their administration and colleagues. The returned tasks were analyzed to answer research question four.

3.3 DATA CODING & ANALYSIS

This section presents rubrics and explains how the data were coded and analyzed to answer the four research questions. So the first section explains how the participants' conceptions of proof was coded and analyzed. Then an explanation is provided for how the participants' solutions are coded and reported. Thirdly, the coding system for how the participants' analysis of student solutions is explained. Finally, a description for how the reasoning-and-proving tasks the participants' selected and or modified were coded is shared. Data from this study is used to explain the coding and analysis process.

3.3.1 Conceptions of Proof

Two perspectives to consider when reviewing secondary teachers conceptions is the nature of proof in mathematics and the handling of it in the classroom (Knuth 2002b). For instance, one may think of the construction of proof as a creative process for mathematicians, but believes students need to be constricted to particular forms based on their novice knowledge of mathematics or believe that students are unable to construct proofs at all. Also, understanding what teachers count as proof is instrumental since it will be the implicit or explicit criteria they use to critique their students' work (Martin & Harel, 1989). The way students' arguments are evaluated will influence students' conceptions of proof.

Table 3.5 was designed to capture the participants evolving conceptions of proof. The themes that were promoted and questioned throughout the course makeup the four main categories: criteria, equity, opportunities, and purpose. Mathematics educators (Ball et al. 2002) and the Common Core State Standards (CCSSM, 2010) specifically promote the equity and opportunity themes. For instance, the standards suggest that the goals are intended for all students in all grade levels. If teachers only believe students of certain ability level are capable of writing proofs, they may exclude them from participation in R&P activities. Furthermore, the course engaged the participants in solving several reasoning-and-proving tasks that were outside geometry, which is the conventional course where students are asked to write proofs.

The other two themes or dimensions in the conceptions of proof table are connected to research on the criteria and purpose of proof. In order for students to develop a clear understanding of what counts as a valid argument, their teachers must hold them accountable and not accept empirical or other non-proof arguments as proof (Bieda, 2010). Additionally, Harel and colleagues have argued that students not be expected to memorize or follow external construction of proofs as is typically the case in classrooms when proof is taught (Harel & Sowder, 1998; Harel & Rabin, 2010). Instead, students should be provided opportunities to think and reason through various proof form, types or representations as teachers provide support. Finally, Bell (1976) explains that proof tasks in schools should also be used to help students understand and learn mathematics. Traditionally proof is only taught in schools to systematize definitions and statements. The single view promotes a distorted conception of the purpose of proof in mathematics.

Table 3.5. Conceptions of proof categories

Conceptions of proof categories	
<u>Criteria: What counts as proof?</u> Argument must show that the conjecture is (or is not) true for all cases. The definitions and claims must be true and accepted by the community. The conclusion follows logically from the argument *<u>A proof may vary along these dimensions:</u>* type of proof; form of the proof; representation used; explanatory power	**<u>Equity: Who should write proofs?</u>** Closed: honors students, Middle: All students can reason, but writing proofs might not be possible for some students Open: All students can write proofs including special educational students
<u>Purpose: Why teach proof?</u> To learn new mathematics To systematize definitions and statements in an axiomatic system To verify truth To communicate knowledge To explain why something is true To explore meaning To construct an empirical theory (Bell, 1976; de Villers, 1990; Hanna, 2000)	**<u>Opportunities: When (how often) should proof be taught?</u>** Not a priority: Time permitting, it comes up in the curriculum or do a little in geometry Special topics or units or courses: will teach proof in geometry and may fit it in here and there in algebra Priority: possible in every unit of every course

The reasoning-and-proving course, through the various course activities, aimed to expand the participants' views of proof beyond conventional conceptions of proof in secondary schools. The interview questions along with the opening course activity and final course reflection explicitly asked the participants their position on the four themes described in table 3.5. These themes address the nature of proof in mathematics and instructional decisions teachers make with regards to implementing reasoning-and-proving tasks.

Starting in the left top corner, the goal would be to recognize teachers developing a full criterion of proof. In the course this was specifically discussed in terms of constructing generic arguments and or demonstrations. The student work analyzed along with the reasoning-and-proving framework introduced the participants to a variety of ways to construct valid arguments.

Moving in a clockwise direction, Knuth (2002b) learned that some secondary teachers do not believe all students should engage in writing proofs. The course took an equitable stance that all students should be provided opportunities to engage in reasoning-and-proving tasks. So statements, which convey a movement away from a closed conventional stance on proving, were coded as evidence of growth in this theme.

The bottom right corner (opportunities) is also a point of contention. If textbooks do not include many opportunities to reason and prove, then how often should R&P tasks be taught? The course espoused a belief that reasoning-and-proving should be integrated into all secondary course units though the explicit engagement in modifying tasks. The traditional view would be that proof is only taught in a few chapters in a geometry course. So talking about enacting reasoning-and-proving tasks outside the conventional geometry course was coded as movement.

Finally, the bottom left theme describes the purposes of proof in mathematics as well as the purposes advocated for school mathematics. The traditional use of proof in school is only to organize definitions and statements into axiomatic systems. The course encouraged multiple solution paths and multiple representations to show the value of proof for mathematical understanding. Additionally, participants were provided opportunities to show why a statement was true, and engaged in communicating arguments orally and in writing within the class community. Coding themes on the purpose of proof identifies movement if the participant suggests opportunities of engagement outside typical two-column axiomatic proofs. The point is that if the participants come to recognize a broad number of purposes, then they may be more likely to implement reasoning-and-proving tasks to support students with attaining the broader set of goals.

3.3.1.1 An explanation of how the data was analyzed to address the four conceptions of proof?

The transcripts from the open-ended questions (part 1) for each of the nine interviews were organized along with the written responses to the course "opening activity" and the final reflection. A table was created for each of the nine participants that resembled table 3.5. Then each of the participant's responses were reviewed and instances that matched one of the four categories were copied and pasted into the individuals table. All information a participant shared during the first interview was bolded in the their table. The data the participants shared during the second interview was italicized and placed into their table. The responses to the third interview and the final course reflection were underlined so as to distinguish when the participants shared their thinking. This process resulted in 46 pages of information into nine tables (one table for each participant).

The data placed into each of the four conception categories were reviewed to identify themes. First, the information the participants explained about their understanding about what is needed for an argument to count as proof was analyzed and categorized against the course criteria of proof as follows:

Course Criteria of Proof

- *Argument must show that the conjecture is (or is not) true for all cases.*
- *The definitions and claims must be true and accepted by the community.*
- *The conclusion follows logically from the argument*

  <u>A proof may vary along these dimensions:</u>

  type of proof; form of the proof; representation used; explanatory power

Secondly, what the participants said about the purpose of proof was also organized into themes across the nine participants. For instance, the participants learned that participants need opportunities to engage in proof tasks to learn what reasoning-and-proving means. Thirdly, the participants' beliefs based on what they said or wrote was grouped to discuss their conception of which students they believe should have access to writing proofs. The final category was analyzed in the same way with respect to the classes the participants believed that proof should be taught and how often during a school year proof should be include. Specific quotes were chosen and shared in the analysis to highlight what a group of participants said and in other situation tables were developed to report the results of what the participants wrote during the opening activity and final reflection and said during the open-ended interview questions.

Growth is explained by a comparison to what the participants said or wrote prior to the course to what he or she shared during, and at the end of the course. For the criteria of proof, all three time periods are analyzed separately. For the purpose, equity, and opportunities category conceptions, there are only two data points: 1) prior to course and 2) interview two and end of the course are grouped together.

The reason the criteria of proof is discussed at all three time points is because the analysis includes the instances in which participants expanded upon and initial characteristic of the criteria of proof. For instance, two participants initially explained that they knew that a proof needed to cover all cases, but after engaging in a particular course activity they explained that their thinking about an argument covering all cases was enhanced. The other three categories were not conducive for the participants to expand upon their initial thinking. The point was to learn if the participants included added new purposes or changed their initial belief as to which students or courses should include proof. When they changed their conception is not relevant, only they a change was detected based on their initial thoughts from the first interview.

3.3.2 Solving Reasoning-and-proving Tasks

The participants solved reasoning-and-proving tasks in the interviews and during the course. The teachers were asked to solve one task in each of the interviews, and solved five problems during the course. The written solutions to all eight tasks for the nine students sum to a total of 71[4] arguments that were be coded and analyzed.

Since the eight tasks vary among mathematics topics and the types of problems could be more or less familiar to some participants, it does not make sense to show improvement over time in their ability to write a proof. Instead the solutions were coded to learn the extent to which each participant was able to construct a proof for each of the eight tasks. Since the tasks are at the secondary level, and data of prospective teachers or practicing teachers' ability to write proofs for this type of content does not exist, this data begins to fill that void.

Many studies on proof employ a five point system for scoring responses including empirical arguments and proofs (e.g. Lannin, 2005; Recio & Godino, 2001; Senk, 1985; A.J. Stylianides & G.J. Stylianides, 2009). This study modified the rubric from the A.J. Stylianides and G.J. Stylianides (2009) study. Their rubric included five argument types, which became the

[4] Karen did not solve one of the course tasks since she was late to class.

main categories for this study. This author then expanded the rubric to include subcategories and a clear and convincing dimension for valid arguments (as shown in table 3.6).

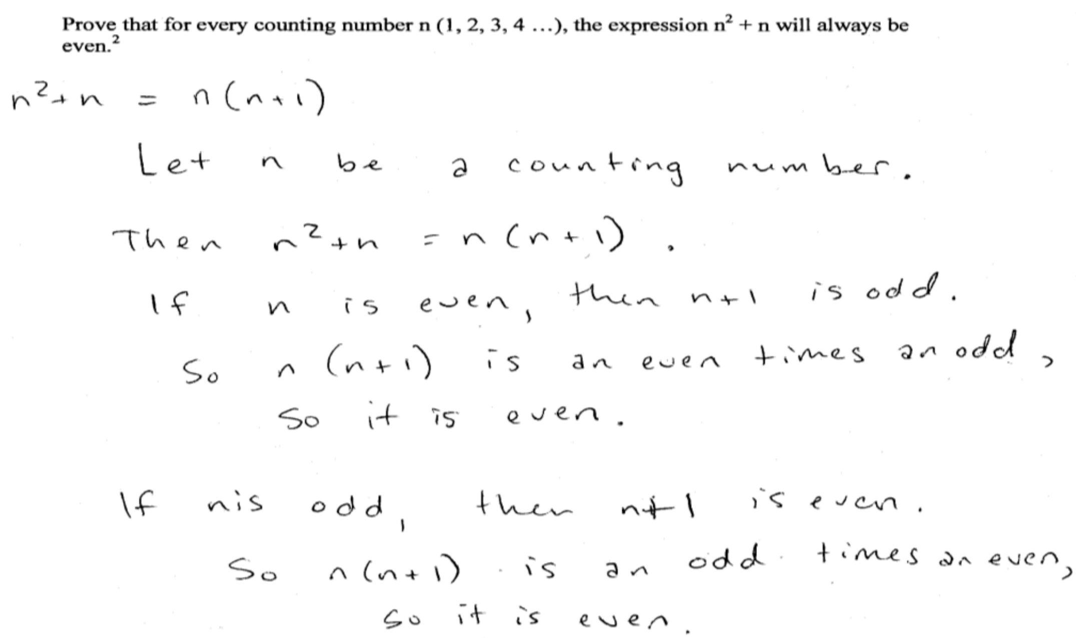
Table 3.6. Reasoning-and-proving task coding tool

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

The first column is the list of main arguments codes that were used in the A.J. Stylianides and G.J. Stylianides (2009) study. The second column shows sub codes that were developed to better distinguish among the various participant solutions. The final column includes directions on how to code a solution that meets the main category code. The three clear and convincing categories are listed below the table and are only applied to valid arguments (A3 or A4).

The code A0 was used on solutions that were incoherent or it was evident the participant did not address the problem situation. A1 was used if the participant was unable to reach a solution or make a generalization. There are two sub-codes to accompany the main codes A0 or A1. The code A2 is applied to solutions where the argument is missing or the argument lacks generality. The various sub codes (1-5) in the second column identify specific issues as to how the solution is limited. The A3 code represents a valid argument, but includes too much interpretation on the part of the reader to count it as proof. In other words, A3 was applied when assumptions were detected in the argument including an assumption about the conjecture the participant is attempting to prove. A4 is a proof, and no sub-codes follow the A4 main codes, which is why the bullets are used, opposed to numbers in the second column. Finally, a plus or minus symbol is used to code all A3 and A4 main codes. A plus is listed for each clear and convincing statement that is represented in a valid argument or a minus is used to indicate that a clear and convincing statement is absent. Therefore, each valid argument code (A3 or A4) is followed by a combination of three plus or minus symbols.

An example is provided to explain how a solution is coded (as shown in figure 3.3). The $N^2 + N$ is always even problem was the first task the participants solved during the first interview. Tanya's solution is mathematically correct, but she includes an assumption when she wrote that an even times and odd is even without justifying why this is true. The valid argument with the assumption means that the solution is coded A3.1. Since the solution is a valid argument, all three clear and convincing statements need to be check. The argument does not include jarring statements, missed defined terms or variables, but there is no clear conclusion. The two cases are addressed without summarizing the argument to explain why the conjecture is indeed true. Therefore, Tanya's solution was coded A3.1 ++ -.



Prove that for every counting number n (1, 2, 3, 4 …), the expression $n^2 + n$ will always be even.[2]

$n^2+n = n(n+1)$

Let n be a counting number.

Then $n^2+n = n(n+1)$.

If n is even, then n+1 is odd.

So n (n+1) is an even times an odd, so it is even.

If nis odd, then n+1 is even.

So n(n+1) is an odd times an even, so it is even.

Figure 3.3. Tanya's solution to the $N^2 + N$ is always even task

In addition to coding all 71 solutions, the challenges the participants encountered while trying to construct arguments were identified across all eight tasks. The challenges provide insight into what prospective teachers may need more support with to construct proofs.

A second coder was trained to account for rater reliability, and coded 18 of the 71 solutions. Agreement was reached on 13 out 18 (72%) main codes, but every researcher labeled valid argument (A3 or A4) was also labeled as such by the second coder. There was one instance in which the second coder labeled a solution as a valid argument and the researcher coded it as an A2. Therefore, four of the disagreements were between A3 and A4 and only one between A3 and A2. There were a total of eight solutions that required a sub-code and there was agreement on seven of the eight possible sub-codes. There were 33 opportunities to include a plus or a minus for the 11 valid arguments, and agreement was met on 26 of the 33 (79%) instances.

3.3.3 Critiquing Student Arguments

The participants critiqued student solutions to tasks they solved in part two of each of the three interviews and were also provided two opportunities during the course to analyze and make judgments about student work. In all, the interviewed participants each analyzed 32 student solutions to five different reasoning-and-proving tasks. The participant responses to the 288 student solutions were coded and analyzed.

Several of the 32 arguments included in the samples were intended to be "distracter" items. These student solutions either fall short of being a proof for some reason or are proofs that do not fit a more traditional view. A total of 12 student solutions are identified as "distracter" items and organized chronologically as to when the participants where asked to analyze the solutions (as shown in table 3.7). These solutions were purposely placed in the set of solutions to create a cognitive conflict to reshape the participants' mental image of what counts as proof.

Each participant's coding of the student solutions were analyzed. Then the 12 identified distracter solutions were further analyzed to determine growth in recognizing various types of arguments. All 32 solutions were coded with the following identifiers: CI & RC (correctly identified & reason correct), CI & RNC (correctly identified & reason Not correct, or IC (incorrectly classified). So in addition to a participant correctly identifying an argument as proof or non-proof they also needed to explain a correct rationale for their choice. While validating the student solutions during the course, the participants were given the opportunity to label a solution as yes (a proof), no (not a proof), or unsure. The participants provided a rationale for any of the three choices. For example, if a participant identified a non-proof argument as such and explained why the solution is not a proof, then he or she was said to have given a correct response. The solutions labeled "unsure" were not marked incorrect since it is believed that saying an argument is indeed a proof when it is not is different than explaining that one is uncertain. In some cases a participant labeled an argument "unsure" and provided a correct explanation for why the solution is or is not a proof. Instead of reporting such a case as correct, it is left as a separate category (unsure) since the participant is still negotiating their understanding of their criteria of proof.

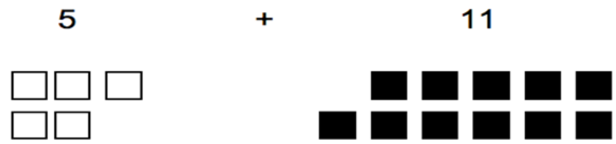
Table 3.7. Twelve Student Solution Distracter Items

Twelve Identified Distracter Student Solutions				
1) N2+N is even	3) O + O = E	4) NxN square window	6) Sticky Gum	8) Calling Plans
Student D: Proof/rationale: Could be considered a rationale since the argument makes assumptions such as: If n is even, then n2 is even.	Student B: Generic example: makes a general claim from a specific diagram example	Student A: Empirical: constructs a generalization from a few examples without justifying.	Student C: Provides Justification; non-proof	Student A: Not a valid argument. An incorrect narrative response. Student B: Proof: counterexample. Provides a narrative general and a specific counterexample.

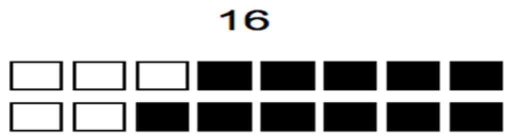
Student E: Proof (generic argument): relies on diagram (specific cases) to generalize for all cases	Student I: Rationale: Correct statement, but assumes too much.	Student B: Proof (generic example) uses a particular case to generalize to any size window.	Student H: Provides justification, but not general to all cases.	Student C: Not valid. Provides a convincing response without attending to the question. Student E: Proof: provides a general counterexample argument.

The results of this analysis are presented in Chapter 4 in a table, such as the one shown in table 3.8. The top row lists the students solutions from A-J. The first column includes the five tasks during the course and interviews that include student solutions. The diamonds (◆) in the cells identify the student solutions within each task that are distractors. The “P” means that the particular argument is a proof and “NP” represents non-proof student solutions. The student B “O + O = E” solution code of 2U:5/9 means that 5 participants correctly identified the argument as proof. The 2U means that two participants said they were unsure. So the remaining two participants claimed the solution is a non-proof. Ratios are shared for all students solutions in which at least one participant misidentified an argument. Finally, in addition to the table, common participant challenges are discussed arcoss all 32 solutions.

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., et al. (2003). Contemporary mathematics in context: A unified approach: Course 3. New York, NY: Glencoe McGraw-Hil

Figure 3.4. Student B solution to the “O + O = E” task

Table 3.8 Tweleve distracter items represted across five student work activities

Proof Codes	A	B	C	D	E	F	G	H	I	J
1) N^2+N is even	P	NP	NP	NP:◆	P:◆					
3) Odd + Odd = Even	P	P:◆ 2U:5/9	NP	P	NP	NP	P	NP	NP:◆	P
4) NXN window	NP:◆	P:◆	P	NP						
6) Sticky Gum	NP	NP	NP:◆	NP	NP	NP	P	NP:◆		
8) Calling Plans	NP:◆	P:◆	P:◆	NP	NP:◆					

Consider figure 3.5, as a model for all possible solutions constructed for any reasoning-and-proof task. The vertical black line separates proof from non-proof arguments. Solutions placed to the far left could be labeled “most definitely not a proof.” Similarly, the solutions to the far right could be labeled "most definitely a proof." On the other hand, just to the left of the black vertical bar are non-proofs, but solutions close to the vertical bar may be labeled proof by some participants. Additionally, just to the right of the black vertical bar, even though these are proofs it is likely that participants would disagree. Over time the goal would be for a community of learners to develop a common criteria of proof so that they would come to agree about the placement of solutions as either proof (right side of black line) or non-proof arguments (left side). The 32 solutions the participants analyzed in this study spanned a spectrum of possible argument types. The CORP design challenge was to select student solutions for teachers to analyze that would foster prospective or practicing teachers understanding of the criteria of proof. Choosing only solutions that are easy to identify would not help teachers distinguish between those close to the vertical black bar, but only choosing student arguments that cause disagreement would not provide a facilitator with information regarding what participants know, which is why some solutions are easier and others more difficult to analyze. Therefore, the results, which will be shared in the subsequent chapter indicate what the participants found to be challenging and what types of arguments were easier to analyze.

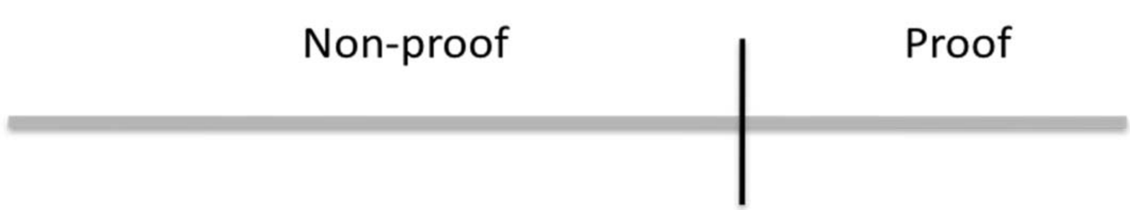


Figure 3.5. Spectrum of solutions to reasoning-and-proving tasks

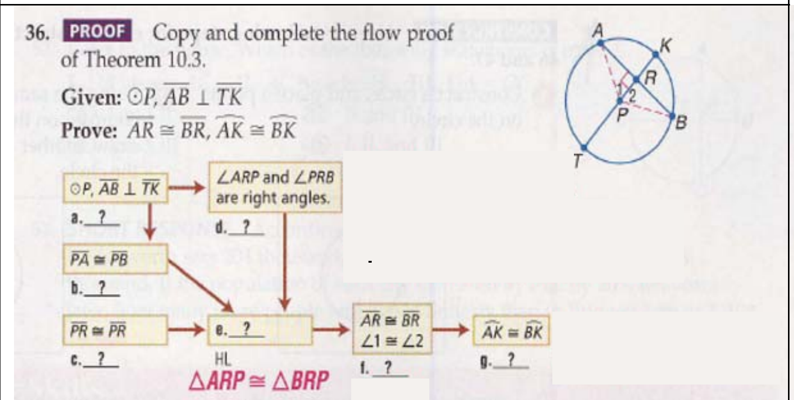
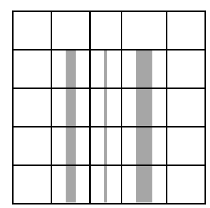
3.3.4 Selecting Reasoning-&-Proving Tasks

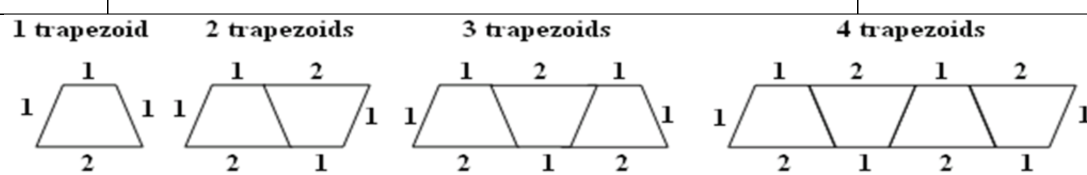
During the course, including the interviews, the participants had the opportunity to identify three tasks that could be used to promote reasoning-and-proving. All the participants selected and modified the two tasks for the two required course assignments. Five participants selected a fifth task, which was discussed during interview three. After the course, two participants identified reasoning-and-proving tasks that they enacted in their classrooms with their students as first year teachers.

A main goal of the course was to prepare participants to select or modify reasoning-and-proving tasks. Modifying tasks includes taking low-level tasks or non-reasoning-and-proving problems and, through the addition or deletion of information, creating high-level tasks that include reasoning-and-proving activities. For instance, if a task requires students to follow a procedure, then questions could be added that ask students to explain why the process works, to make connections using context, or to link several mathematical representations to develop a deeper understanding of the method. Given, that the participants were asked to select reasoning-and-proving tasks beyond just any high-level mathematics problem, the identified task needed to prompt students to make a mathematical generalization and or develop an argument to support a conjecture. Selecting and or modifying tasks that allow for students to develop a generic argument could support students in developing a proof (Lannin, 2005; G. Stylianides, 2010).

The rubric used to analyze the tasks is a combination of the reasoning-and-proving framework and the task analysis guide (TAG). With regards to the TAG, instead of parsing the tasks into memorizing or procedures without connections, any activity that fits into either one of these categories were labeled low-level. Additionally, the procedures with connections and doing mathematics problems were grouped into a single high-level category. The two broad mathematical components in the R&P framework, making generalizations and developing arguments, serve as the second dimension. Furthermore, some tasks neither provided students an opportunity to make a generalization or provide, which are labeled either low or high-level non-reasoning-and-proving tasks. Therefore the six possible codes in order from lowest cognitive demand to highest level cognitive demand with reasoning-and-proving are: L non-R&P (Low-level non-reasoning-and-proving), LG (low-level make a generalization), LP (low-level provide an argument), H non-R&P (high-level non-reasoning-and-proving), H-L P (high-level make a generalization), and HP (high-level provide an argument). Four example tasks are provided in table 3.9 to serve as a guide for coding teacher-selected tasks that are reasoning-and-proving types.

TABLE 3.9 EXAMPLES OF THE FOUR DIFFERENT CATEGORIES OF PROOF TASKS

Cognitive Level of Tasks	**Reasoning-and-Proving Activities**	
	Making Mathematical Generalizations	**Providing Support to Mathematical Claims**
Low Level	**MAKING CONJECTURES Complete the conjecture based on the pattern you observe in the specific cases.** **29. Conjecture:** The sum of any two odd numbers is __?__. $1 + 1 = 2$ $7 + 11 = 18$ $1 + 3 = 4$ $13 + 19 = 32$ $3 + 5 = 8$ $201 + 305 = 506$ **30. Conjecture:** The product of any two odd numbers is __?__. $1 \times 1 = 1$ $7 \times 11 = 77$ $1 \times 3 = 3$ $13 \times 19 = 247$ $3 \times 5 = 15$ $201 \times 305 = 61{,}305$	
High Level	Refer to the trapezoid pattern below. 1.What is the perimeter of the pattern containing 12 trapezoids? 2. Use the diagram to describe how you can find the perimeter of a pattern containing any number of trapezoids. 3. Find a second way to find the perimeter of a pattern containing any number of trapezoids.	Jordan and Adam decided to prove their conjecture (that the sum of two consecutive numbers was equal to the difference between the squares of the two consecutive integers) by drawing a picture as shown below. The bell rang before they could label or explain the picture so they quickly wrote, "the white squares tell the story", and put it in the pile on the teachers' desk.  What do you think Jordan and Adam were trying to communicate with this picture? Does this picture constitute a proof? Why or why not?



All of the participants selected two tasks (2 times 9 or 18) and modified the same two tasks (2 times 9 or 18 more) during the course (36 total) and five participants brought a task to the final interview (5 additional). Two teachers selected and implemented five tasks each during their first year as a classroom teacher (10 additional). So there are a total of 51 (36 + 5 + 10= 51) tasks. In addition to coding all 51 tasks, the initially selected 18 course tasks (2 per each participant) were analyzed against the modified version of those tasks. The five interview three tasks are discussed separately as are the ten participant "classroom" implemented reasoning-and-proving tasks.

In order to report reliability coding for the selected and or modified reasoning-and-proving tasks, 12 were randomly selected and coded by a second coder. Ten of the 12 tasks were classified the same for an 83% reliability rating.

4.0 CHAPTER 4. RESULTS

The results of the analysis as described in chapter three organized by each research question are presented in this chapter. Specifically, section 4.1 details the participants' conceptions of proof. Data collected during individual interviews conducted outside the regular class meetings and written responses to two in class prompts are used to provide evidence. Section 4.2 reveals the results of the participants' abilities to construct valid arguments, which included five opportunities during the course, as well as opportunities presented during each of the three interviews. All responses were coded using the reasoning-and-proving coding tool. The eight tasks are analyzed and the participant responses presented. Finally, an analysis of the participants' arguments across the eight tasks is reported to uncover and identify changes with respect to their solution methods over time.

Section 4.3 addresses the participants' skills in distinguishing between proof and non-proof arguments. The participants analyzed two sets of student solutions during the course and one set of student work during each of the three interviews. In addition to identifying a solution as proof or non-proof the participants' reasons for making a decision are considered and reported. The final research question reported in section 4.4 showcases the results of the participants' abilities to select and or modify reasoning-and-proving tasks. All of the participants selected and modified two tasks during course, were asked to bring one to the final interview, and were invited to share tasks they enacted as first year teachers. The tasks were analyzed along two dimensions: 1) level of cognitive demand (high or low) and 2) extent to which the tasks includes reasoning-and-proving activities (make a generalization, provide an argument, or not a reasoning-and-proving task).

4.1 PRE-SERVICE TEACHERS CONCEPTION OF PROOF

The results in this section are in response to the first research question:

> *1. How do pre-service teachers' conceptions (i.e. purpose of proof, what counts, proof in secondary courses) of proof change over the duration of a course focused on reasoning-and-proving?*

The participants' conception of proof is assessed according to four main categories (as shown in table 4.1). Two of the perspectives pertain to a mathematical understanding of proof for teaching (1st column of table 4.1): a teacher's criteria of proof and the purposes for teaching reasoning-and-proving activities. The second pair of conceptions (2nd column of table 4.1) address the participants' perspective about which students they believe are capable of engaging in reasoning-and-proving activities, including the courses that should contain it, and how often students in such courses should be provided opportunities to reason-and-prove. Each participant answered questions during interviews and wrote written responses to two course prompts, which relate to each of the four categories. The analysis was organized along each conception of proof category and summarized to explain changes in the participants' thinking.

At the beginning of the first class session, the participants wrote individual responses to a series of questions about proof. These written responses, along with the open-ended participant replies during the first interview, are combined to establish each participant's initial conception of proof. The second data point for establishing participants' conceptions of proof is based solely on the responses to the open-ended questions during the second interview. The third and final data collection consists of the questions posed during the third interview and the final written class reflection, which included two questions about proof. The participants' conceptions of proof over the three time periods across four dimensions are reported according to criteria, purpose, equity, and opportunities.

Table 4.1. Four conceptions of proof categories

<table>
<tr><th colspan="2">Conceptual R&P Themes</th></tr>
<tr><td><u>Criteria: What counts as proof?</u>
Argument must show that the conjecture is (or is not) true for all cases.
The definitions and claims must be true and accepted by the community.
The conclusion follows logically from the argument
<u>A proof may vary along these dimensions:</u>
type of proof, form of the proof, representation used, explanatory power</td><td><u>Equity: Who should write proofs?</u>
Closed: honors students,
Middle: All students can reason, but writing proofs might not be possible for some students
Open: All students can write proofs including special educational students</td></tr>
<tr><td><u>Purpose: Why teach proof?</u>
• To learn new mathematics
• To systematize definitions and statements in an axiomatic system
• To verify truth
• To communicate knowledge
• To explain why something is true
• To explore meaning
• To construct an empirical theory
(Bell, 1976; de Villers, 1990; Hanna, 2000)</td><td><u>Opportunities: When (how often) should proof be taught?</u>
• Not a priority: Time permitting, it comes up in the curriculum or do a little in geometry
• Special topics or units or courses: will teach proof in geometry and may fit it in here and there in algebra
• Priority: possible in every unit of every course</td></tr>
</table>

4.1.1 Criteria of Proof: What is proof and what counts as proof?

The course criterion was used as a measure for judging the quality of each participant's conception of proof (as shown in table 4.2). The criterion of proof from the course is parsed along the second row of table 4.2. The term *sensible argument* was used instead of *logical argument* since "logical" has a specific meaning in proof writing. The participants used the word *logic* and *logical argument* to mean *sensible and mathematically correct*, opposed to various logical (e.g. contradiction, contra-positive, direct) or illogical (e.g. converse) forms. Participant names are listed along the first column of the table. The numbers in the cells represent the time period (1: prior to start of course, 2: between the 4^{th} and 5^{th} class meeting, 3: at the conclusion of the course) when each participant described a particular characteristic of the criteria. The bolded numbers mean that the participant expanded upon what he or she said in the previous interview. The cells that include multiple numbers represent the case where a participant mentioned the characteristic multiple times. Blank cells indicate that the participant did not discuss the particular criterion factor. For instance, Tina has a '1, 2, 3' in the "true for all cases" cell, means that during each of the three time periods Tina explained that a proof must cover all cases, but did not expand on her initial thoughts about this characteristic. Thus, the '2, 3' code in Tina's row under the "claims and statements accepted by community" means that she did not mention this characteristic prior to the second interview. Furthermore, Tina has a '1, **2**, 3' code in the "type, form, and representation can vary," which signifies that she mentioned the possibility of multiple forms or representations during the first interview then expanded upon what she said in the second (bolded 2), and referenced the variety of forms proofs can assume in the final interview without explaining any new understanding from the second interview at the conclusion of the course.

Table 4.2. Participants' criteria of proof compared against course criteria

Participants' Criteria of Proof							
	True for all cases	**Counter-example**	**Claims and statements accepted by community**	**Sensible argument**	**Conclusion should be included**	**Clearly articulated language**	**Type, form, and representation can vary**
Nathaniel	1, 2	1	2	1	1	1	1, 3
Tanya	2		2	1, 2, 3			1
Karen	1, 2, 3		2	3	1		2, 3
Tina	1, 2, 3	1	2, 3	1, 2, 3		3	1, 2, 3
Lucy	1, 2			2			2
Uma	1, 2, 3		2, 3	1, 2, 3		1	1, 2
Brittany	2		2, 3	1, 2, 3	1, 3	3	1, 2, 3
Katie	2, 3		2, 3	2, 3	1, 3	1, 2, 3	1, 3
Katherine[5]	1, 2, 3	2	2, 3	1, 2, 3	2		1, 2, 3

[5] Katherine's first interview data was lost so the information is only based on what she wrote during the first class.

Many participants claimed their understanding of proof did not change. However, one finding is that no component was mentioned by all of the participants prior to the start of the course and every characteristic was added by at least one participant during a subsequent interview (as shown in table 4.2). For instance, three participants (Tanya, Brittany, and Katie) did not mention that proofs need to be true for all cases prior to the start of the course, but they each added the requirement to their criteria during the second interview. Looking at each column there is at least one ‘2’ or ‘3’ without a ‘1.’

Also every participant expanded his or her criteria of proof by at least two new characteristics, and four participants (Katherine, Katie, Brittany, Karen) expanded upon or described four new components after the course started. Looking across any row there is at least one ‘2’ or ‘3’ without a ‘1,’ and most (7 out of 9) participants expanded on (a bolded number) a previously discussed characteristic. For example, Katie never mentioned the characteristics generalize for all cases, that a community must agree upon what is acceptable, or that true mathematical statements need to be organized into a sensible argument prior to the course until the second interview. Then during the third interview she explained how her view of the form of proofs changed from a more formal structure to where she now believes a proof could assume many forms. Therefore, even though only Nathaniel discussed all seven characteristics, all nine participants expanded upon their initial criteria of proof.

To better understand how the participants thinking about the characteristics expanded, the next few sections will explain three of the four most discussed criteria and then discuss the three least mentioned components as a group. The second and fifth columns are similar in that every participant discussed both, and two-thirds of participants (6 out of 9) first started talking about each characteristic during the first interview. The difference is that two participants expanded upon the "must be true for all cases" criteria, and no participant spoke specifically about how their understanding of a "sensible argument" changed. Therefore, columns two (must be true for all cases), four (claims and statements accepted by community**)** and eight (type, form, and representation can vary) are explained to gain a deeper insight into what the participants said and how they changed their understanding of these three characteristics. Finally, an explanation is provided for the three criteria (columns 3, 6, 7) that were not mentioned by many participants.

4.1.1.1 Proof must be true for all cases

The fact that a proof must cover all cases was a criterion discussed by all nine participants (as shown in table 4.2). Two-thirds of the participants (6 of 9) mentioned the need to cover all cases prior to the start of the course. Many participants made general comments about the need for proof to cover all cases prior to the start of the course and a few did not make mention of it until the second interview. For instance, Lucy said, "If it's like a written proof, then it should prove something is true 100 percent of the time for every case." However, just over half of the participants (5 of 9) made specific reference to a clearer understanding of covering all cases during the second interview. The sequence of three tasks specifically reshaped their conception of the use of examples to generate a generalization. They realized that allowing themselves and students to write a formula from a few examples is not a secure method of proof. This new realization changed their thinking about how one needs to show a situation is true for all cases. For example, Karen talked about how the sequence of three tasks during the first class changed her own thinking about the use of examples and showing a situation is always true:

> *I really liked doing the problems of squares, problems like the 60 by 60, and then the dots on the circle and then the – I mean the counterexample is pretty crazy, but I kind of like doing that because I have always been used to finding a pattern, make a conjecture all that, and then we did that circle thing and we saw that it doesn't always hold. I think that's been pretty interesting to just kind of remember that just because it works for a few*

cases, a few situations, it doesn't mean that it's going to hold forever. So that was pretty – a cool thing to remember and definitely something that I want to try to use next year when I try to do these sorts of things.

Prior to the course, Karen knew a few cases were insufficient for proof, but she admitted that she accepted generalizations based on a few cases. Therefore, the course supported her change in understanding that it is not acceptable to assume the truth of a generalization from a set of examples. Karen's comment was representative of just over half the participants (5 of 9).

4.1.1.2 Claims and statements accepted by community

A shift in thinking was evident in understanding that proof is a communal activity (fourth column in table 4.2). Prior to the course, no participant mentioned that an argument needs to be accepted by others. In other words, they came to understand that proof is not an individual activity, and the classroom (or mathematical) community must accept the claims and statements used in an argument. Eight of the nine participants mentioned this new realization during the second interview. Katie explained that prior knowledge and the use of it in a community as "the toolbox of statements" that are acceptable. She made the follow comment:

For an argument to count as a proof, it needs to be logical and clear. You have to have like a set of understanding for whatever community you are kind of working with, and that can be different you know for each community, it has to be, you have to have a sort of a set of statements, what I consider to be like in your little mathematical toolbox.

Katie references the importance of a community agreeing upon the truth of mathematical statements. These truths within a community is accepted as prior knowledge or what Katie and others called their 'mathematical toolbox.' These community wide accepted truths can then be inserted into arguments so that the statement does not need to be justified again. Tina explains her understanding of the acceptance of claims in a proof as:

Like what can we allow to be prior knowledge kind of seems like it's not really something that an outsider can determine. It's more something that you have to be the actual teacher to make a judgment call on whether that's allowed. Yeah. Just thinking about assumptions and things that they can make about numbers. As the teacher, you're kind of the only one that knows what you've done all year and what they can use without explaining or proving necessarily inside their proof.

Tina is more specific about the inclusion of assumptions and how a classroom teacher would need to make clear what is accepted and what students would need to further explain. These teacher decisions would be based on what the classroom community previously proved.

4.1.1.3 Type, form, and representation of a proof can vary

The participants expanded their view of proof throughout the course with respect to the various types, possible representations, and forms (last column of table 4.2). Prior to the course, most of the participants already knew that proofs could take on different forms, but it was limited to two-column or paragraph. The last column in table 4.2 shows a '1' for 7 of the 9 participants, meaning that they knew proofs could assume a variety of forms and representations, but during the subsequent interviews most (8 of 9) participants commented on how the course changed their view as to the extent in which a proof needs to be formal. The eight participants implicitly or explicitly indicated that after engaging in various course activities such as analyzing student work, reading the article about the reasoning-and-proving framework, and developing the criteria of proof, their view of what counts as proof expanded to accept that proofs do not necessarily need to be formal and can include diagrams and or narrative language.

For example, during the third interview, Nathaniel explained his new understanding:

I guess before coming to this course I said I really hadn't had much experience teaching what I would've called reasoning and proving tasks, when I initially heard reasoning and proving, I would've thought of pretty formalized proofs. But after kind of learning a little bit more with them trying to say about the process of going through like the reasoning *and developing of the conjectures and then finalizing it out with actually like mathematically like proving your arguments.*

Here, Nathaniel discusses his understanding of reasoning-and-proving as a set of activities that lead toward proof. During the first interview, Nathaniel mentioned proofs could take on various forms, which is represented by the '1' in the last column. In the quote above, Nathaniel explains how the course changed his thinking about his criteria of proof; specifically how it does not need to be a formal argument. This change is articulated during the third interview and represented in the table with the '**3**.' He spoke implicitly about the reasoning-and-proving framework in that there is a range of activities that teachers can engage students in doing prior to writing a formal proof that could include narrative language or diagrams.

Uma is an interesting case since she mentioned multiple forms and representations during her first interview, but she commented more specifically during the second interview on the struggle she was having about the inclusion of diagrams in a proof. She explains:

I mean, I guess I'm – when I see an argument that's just based on pictures, I'm always a little leery about them. But I feel like if they – if they use the pictures and explain something, some people are more visual learner, so I do think the pictures are acceptable. We have been arguing about that a lot in class. But you just need to make sure that they're clear about what they're drawing.

She clarifies that she is leery of pictures alone, but is now reconsidering diagrams if they are clear and in conjunction with an explanation. This change in thinking shows how Uma expanded her conception about acceptable proof representations.

Finally, Katie seemed to posses a full conception of the various forms, types, and representations at the beginning of the course, but she too admitted that the course clarified her thinking about what is an acceptable proof. During the first interview she said:

Yeah, a full proof could sort of be anything, I think it could be a paragraph, I think it could be a picture, I think it could be… anything, yeah, I don't think that there's a single rule for what a proof should look like. It can use numbers, it can use symbols, it can use words and pictures, pretty much anything.

During the third interview, Katie indicated that she experienced growth with respect to what a proof could be. She explained:

Yeah, I think that if anything changed it just it gives me a clearer picture of what it involves. For example I know that we can have a paragraph it could be a picture, you know with some words, it could be, you know it doesn't always have to contain like a generalized statement like a function or…Yeah, and just like the makeup of it
Yeah, I really had just more of the formal idea, like the two column proof and the proofs we did in college which were more like by induction or just the, in the one article we read the proofs, not as much the proofs that explain, so, you know this class maybe enlightened me on the proofs that you can do in high school that are different from the ones we're asked to do in college.

Even though Katie initially portrayed a broad view about the style of a proof, she still had a formal conception of how words and symbols could be arranged in an argument. She explains "the makeup of it" has changed for her; where the information contained in a proof should convey understanding opposed to following a particular form. The way students choose to demonstrate understanding could include words and does not necessarily need to include symbols. Katie seemed to have a complete conception of logical proof forms, or what they could look like or contain, but was able to the recognize the limitations of her initial view during her final interview.

4.1.1.4 Summary of participant changes with respect to criteria of proof

All of the participants expanded or changed their conception of the criteria of proof. Mostly the participants showed growth along three measures: 1) reconsidered the form, type, and representation of a proof; 2) learned how relying on specific cases to make a generalization is not a secure method for constructing a proof; and 3) statements and claims made in an argument must be agreed upon by the classroom community. The final point, addressed in this section, is whether or not the participants articulated a complete understanding of the criteria of proof.

As noted in table 4.1, this section addresses "what counts as proof." The seven characteristics listed across the top of table 4.2 align with the question, but they are not of equal weight. For instance, not all participants commented on a counterexample. Some participants, using their prior knowledge, explained that a statement must be true for all cases or one must find a counterexample to provide evidence as to why a statement is false. However, this was not a learning objective in the course. Additionally, the need for clearly articulated language and summarizing an argument (columns six and seven in table 4.2) was discussed, but neither was addressed as essential components to accepting an argument as proof during the course. In other words, focusing on concise language was secondary to making sure the participants understood that a valid mathematical argument must be true for all cases.

While just over half of the participants (5 of 9) discussed six or more of the characteristics, not all of the participants were able to articulate the same level of understanding. For example, Katherine addressed six of the seven characteristics of proof across the three time periods and at the end of the course summarized her understanding in the following way:

> *Mathematical proof is showing that a conjecture is true or not true for all possible cases. In order to form a proof, the student must generalize (or be supplied) a pattern, a conjecture, and an argument. There are two types of proof arguments: generic arguments, which generalize an example and demonstrations, which show a complete generalization. Proofs can take many different forms, including diagrams, paragraph proofs, and 2-column proofs. In order for a proof to be valid, it cannot use any statement, which have not previously been proven.*

In her first sentence, Katherine addresses the true for all cases and the possibility of a counterexample (not true). Secondly, she shares her understanding of the reasoning-and-proving framework with her explanation of reasoning as looking for patterns and making a conjecture prior to making an argument. Then she unpacks the term “proof argument” to include the two types addressed in the course: generic argument and demonstration. She also includes various proof forms. Finally, Katherine adds the importance that all mathematical claims made throughout the argument need to be accepted by the community. While Katherine did not list all seven characteristics, she addressed specific course topics to explain her understanding of what counts as proof at the end of the course.

On the other hand, Lucy struggled the most to articulate a broad understanding of the criteria of proof. She consistently shared one-sentence statements and either repeated her sentences in subsequent interviews or just said that her definition did not change. For instance, Lucy started with a sentence for her understanding of proof by saying, “a proof means that something is true 100% of the time for every case.” In follow-up interviews, Lucy stated her thinking about proof did not change, but added that a proof could assume many different forms during the second interview. So even though Lucy was unable to articulate a complete criterion of proof, she did expand on her original conception.

Therefore, eight of the nine participants were able to articulate a broad conception of the criteria proof based on what was addressed in the course by the last interview. All of the participants expanded their view by at least one characteristic. More specifically, the participants expanded their view of proof beyond a formal object to understand that proof is part of a set of activities that starts with making generalizations. Several participants also noted that the sequence of three tasks completed during the first class changed their thinking about the uses of specific cases to generalize a mathematical situation. Finally, most participants articulated that proof is a communal activity, such that mathematical claims and statements made in an argument need to be accepted truths. If the statements are not accepted, then they also need to be proven within the larger argument. Overall, the participants' criteria of proof changed throughout their engagement in the course and most were able to articulate a complete understanding of what is needed for an argument to qualify as proof.

4.1.2 Purpose: Why teach proof?

The rationale behind the question "why teach proof" is that if teachers have a broad view of the purpose of proof, then teachers will find a variety ways to incorporate it into the curricula or see it as an important practice to address. Hanna (2000) recommends that proof in mathematics classrooms should focus on explanation: "The fundamental question proof must address [in the classroom] is 'why?' (p. 8)." While asking 'why' is useful in supporting students to explore mathematical ideas, this research does not claim that a specific purpose should take priority over others. This section will report on the purposes the participants shared prior to the course and then again during and after the course to learn if their conception of the purpose of proof expanded. Finally, the purposes that the participants identify will be compared to those researchers suggest for school mathematics to see the extent to which the participants gained a broad conception for why they will teach proof and how their views align with those of researchers.

4.1.2.1 Purpose of proof before the start of the course

The participants were not directly asked about the purpose of proof in secondary mathematics courses, but the following two questions contributed to the participants' communication of their thoughts on the issue. The first question was asked during the first interview and the second was part of the course opening activity in which they wrote out their answer. The opening activity was the first prompt the participants were given during the first class meeting. The responses to these questions along with other comments the participants said during the first interview and opening activity are used to gain their initial perception of the purpose of proof.

1. What role do you think proof should play in the secondary mathematics classroom?
2. Is it important to engage secondary mathematics students in proof-related activities? Why or why not?

Three themes capture what the participants believed to be the purposes of proof before the start of the course that include: 1) the organization of definitions and mathematical statements; 2) gain a deeper understanding about the truth of mathematics concepts; and 3) develop logical and rationale thinking skills. No one participant communicated all three of these purposes, but each of them mentioned at least one of the three purposes and no one discussed roles of proof beyond the three.

Three participants (Tina, Uma, Katie) explained axiomatic structures and learning to organize theorems and definitions as a purpose to engaging students in proof. The short explanations the three participants shared varied, but their responses seem connected to what the three participants believe is a proof. For instance, Katie said:

> *Begin with basic, universally accepted concepts (axioms) and reason with them to arrive at the desired result. Proof requires substantial justification of an argument.*
> *You've proved a mathematical statement when each "step" contains no assumptions. Each "step" is justified using universally accepted axioms, and the result is clear*

Katie's statement is a longer version and more directed at what she believes is a proof than what Tina and Uma shared. Uma's thoughts are related and connect to another purpose (develop logical and rationale thinking skills) when she said:

It is important for students to develop an organized and thought provoking way of thinking and be able to provide valid arguments. In addition, it is important that students understand theorems and rules that they use and why they work.

Therefore, at the beginning of the course these three participants have identified the need for students to learn proof so that understand theorems and how to organize them in logical steps.

Six participants (Nathaniel, Tanya, Karen, Tina, Lucy, Brittany) believe that proofs are important since they promote a deeper understanding of mathematical ideas. When students understand the truth of theorems or why procedures work, they will build a stronger mathematical foundation, which leads to better recall or reconstruction of knowledge to solve new problems. Brittany, Karen and Tanya each shared examples to convey their thinking about the benefits of teaching proof. Brittany and Karen's views are similar and are examples of writing proofs to better understand the concepts. Tanya goes one step further to say that after they prove a theorem they will not only know it; they will be better position to reconstruct it if necessary. Karen and Tanya's comments are shared:

I think it should play a much bigger role than when I was in school, because when I was working through all of those things with the Pythagorean theorem, I never knew where it came from. But working through the proofs for it - at least the things that we're kinda like proofs for my kids to work through, I figured out where the heck it started. So I think it would give - by thinking about why it holds true always and why something is what it is, will help give the students a deeper understanding of that idea. (Karen)

It is important because it helps students develop a deeper understanding of mathematics if they can prove it rather than just being told that something is true. For example, students might forget that an odd times an odd is always odd. But if they prove the fact, then they will be more likely to remember it because they understand why it is true. If they forget, they can reconstruct a proof to remind themselves of this fact. (Tanya)

Three participants (Katherine, Uma, Lucy) discussed the final purpose (develop logical and rationale thinking skills) as to why students should write proofs prior to the start of the course. The communication around this reason did not spark examples or experiences. Instead the participants stated that learning to think logically is an important life skill and writing proofs is an activity that would provide noteworthy experiences in developing the skill.

Therefore, most participants believed that proofs should be taught in secondary classrooms so that students have the opportunity to gain a deeper understanding of mathematical concepts. Only three participants (Tina, Uma, Lucy) identified more than one purpose of proof and three others created an example or during on personal experiences for why students should learn to construct proofs.

4.1.2.2 Purpose of proof during and after the conclusion of the course

As a group, during the follow-up interviews the participants explained new purposes for why they believe secondary students should engage in writing proofs. The three purposes that were discussed prior to the course continued to be important, but most of the participants shared four new reasons for including proof activities in the secondary mathematics curricula: 1) learn what is proof, 2) communicate mathematical truth, 3) build ownership or authority of the content, and 4) develop an ability to construct a proof.

The sequence of three tasks (Squares, Circle & Spots, Monstrous counterexample) the participants solved during the first class and reading the case of Nancy Edwards (class 5) caused some participants to realize that students may believe that examples are enough evidence to count as proof. Prior to the course no participant commented on the need to enact proof activities with the purpose of supporting students with learning what is proof. Six participants (Tanya, Karen, Lucy, Uma, Brittany, Katherine) focused on the idea that they will need to scaffold students learning of proof in particular that any number of examples is not enough. Lucy discussed this new realization during the second interview:

> *I think before I started this class, I kind of took it for granted that students would know when something was completely – when something was proven, and now I'm kind of seeing that they don't – they kind of have to be taught. Like, yeah, just because you proved it for five cases doesn't mean it's true for every case.*

Lucy shared a similar comment during her third interview.

During her third interview, Karen connected this new purpose of writing proofs with what she said during the first interview about enacting proof tasks so that students come to understand mathematics while making a connection to the case of Nancy Edwards:

> *You can't just assume that they're going to know examples aren't enough. That kind of goes along the line with things that you need to work on from the start. I don't know whose case it was, but those cases helped me to see that, too, and to take tasks maybe not always just because there's a big mathematical idea but I think somebody picked a task where the math was a little bit simpler, but the point was to figure out how do you prove your answer. That was a good thing to understand*

Tanya commented on this same purpose and connected her new thinking to the series of three tasks during her second interview:

> *Just the challenge of how you get students to be convinced that they're, that they have a proof. And that they haven't just made an argument, like how do you teach students that it's a good proof?*
> *I think that activity that we did on the first day was helpful in showing the 30 septillion whatever, that counterexample was useful of just in thinking about how wanted to make sure students understand that they have to check every possible case.*

Therefore, these six participants who spoke about the importance of engaging students in reasoning-and-proving activities to support students with learning a criteria of proof seemed to have made this realization based on course activities.

Three participants (Tina, Tanya, Katie) added the purpose of proof to include the communication of mathematical ideas. None of the three participants referenced course activities, but they suggested that proof could develop students' ability to articulate their thoughts. For instance, during the third interview Katie said:

> *R&P plays a big role with developing classroom norms & math-talk learning communities. Good opportunities to engage in R&P lead to norms like communication, authority, etc. that we want students to have.*

Katie's comment leads into the final new purpose for including proof in secondary classrooms: develop student's mathematical authority.

As with improving communication, just under half of the participants (4 of 9) (Katherine, Nathaniel, Katie, Tanya) mentioned mathematical authority without making specific reference to a course activity. Tanya shared comments about this during the third interview. She explained a connection with supporting students to reason-and-prove and how that process will lead to gaining mathematical knowledge. As students are provided opportunities to communicate their understanding of content, they will develop ownership. Tanya said:

> *I think that proof is connected to helping students develop mathematical authority. Students cannot have authority if they honestly cannot say whether or not their solution is correct. I think that helping students develop reasoning and proving skills will help students gain confidence in their solution and demonstrate more mathematical authority.*

Therefore, it is difficult to know exactly how the course may have influenced their thinking of mathematical authority, but several participants recognized it as a reason for implementing reasoning-and-proving tasks.

Even though each of the participants explained at least one new purpose for proof in secondary mathematics, most continued to discuss what they shared during the first interview. Since the course was not trying to downplay any purposes of proof, it is encouraging to report that participants continued to view their original reasons for proof where these new purposes simply expanded their view of the purpose of proof.

Finally, an interesting finding is that two-thirds of the participants (6 out 9) articulated a broader view of proof as a set of activities. In other words, during the first interview Katie shared a purpose of proof in which to prepare students to organize axioms, definitions and mathematical statements in a step-by-step order until the result is reached. During the second interview, she explained that this is how she viewed the purpose of proof as an undergraduate mathematics student. Since her view of proof expanded to include reasoning-and-proving, Katie as well as five other participants (Nathaniel, Karen, Tina, Lucy, Uma), began to talk about the need to engage students in the spectrum of activities so that students come to understand how to construct a proof. Katie realized that as a teacher she does not want to just focus on the end result of proof as an object. She explained that all students in all courses need opportunities to reason and develop their own conjectures just as she did in the course. During the second interview Katie explains her thinking:

> *Just whenever I was in, you know whenever I was an undergrad I did a lot of proving, but I honestly didn't think much about it in terms of…I guess I definitely made sure that my proofs were valid but I didn't think about the different parts of it you know what I need to do, I guess just thinking about it in terms of teaching it, I'm just getting a little different viewpoint of what it is. Before you know I knew it had to be logical and you know I had the axioms that I could use, but I guess just thinking about it from the other side is definitely just giving me a clearer picture.*
> *Just basically in any class because it's not just about the proof, it's about the logical reasoning and you know being able to build an argument and defend it and I think that's so applicable in real work situations too so I think so far this class has shown me the value in taking the time to do it.*

The five other participants were not as reflective as Katie, but they too recognized the course provided them with a fuller understanding of the need to provide student with a variety of tasks so that the focus in not only on developing a complete valid argument. Students need different reasoning opportunities to support them in developing a proof. During the third interview Nathaniel shared his detailed perspective of his revised thinking when he said:

Since the last time I think I've had a more of a developed understanding of what they mean by reasoning, as like, leading up to the proof, and like the initial thinking where the student needs to identify a pattern or formulate some conjecture, and the process the student can engage in there, to maybe give them a more deep mathematical understanding they can then lead and assist in helping them to come to a more formalized proof, so that'd probably be the main change of my thinking of both reasoning and proving together.

Therefore, the majority of participants were able to articulate that a purpose of reasoning-and-proving is to support students with constructing formal deductive arguments over time.

Overall the participants articulated a total of seven purposes for engaging secondary students in reasoning-and-proving activities. Over the past several decades, researchers have expanded their list for the purpose of proof in secondary mathematics to also include seven purposes:

Researchers lists of purposes of proof

1) To learn new mathematics
2) To systematize definitions and statements in an axiomatic system
3) To verify truth
4) To communicate knowledge
5) To explain why something is true
6) To explore meaning
7) To construct an empirical theory

(Bell, 1976; de Villers, 1990; Hanna, 2000)

The participants identified several of the purposes in the researcher list. However, two noticeable reasons to engage students in writing proofs that garnered attention from the participants during and after the course are not in the researcher list. The course promoted the importance of implementing reasoning-and-proving tasks so that students come to understand a common criterion of proof. Additionally, the participants explained that reasoning-and-proving tasks could be implemented over time to support students with developing a proof that meets the designated criteria. Therefore, the researcher purpose highlights a variety of goals for implementing a proof task, but the participants identified a couple fundamental reasons not on the research list (to learn what is a proof, reasoning activities to support students with producing a proof) that may be useful for secondary teachers with supporting students to construct proofs.

4.1.3 Equity: Who should write proofs?

The nine prospective teachers who participated in this study will have classrooms of their own and will need to make decisions about the opportunities they provide their students. If they understand what reasoning and proving is and the purpose of teaching secondary students proof, then it is more likely they will make choices to include reasoning-and-proving tasks in the course they teach. The next question then is which students should be provided such opportunities. Previous research suggests that high school mathematics teachers believe that only honors or high achieving students should have access (Knuth, 2002b). However, the Common Core Mathematics Standards (CCSSM, 2010) recommends that all students regardless of their ability should have access to constructing and critiquing arguments across all content. Therefore, this section will report on what these participants believe about reasoning-and-proving as an equity issue.

In general, most participants believed that all students should have access to reasoning-and-proving while at the same time some hinted at reservations for including it in their future classrooms. Some participants changed their view on the topic over time, since they did not think much about proof as an element in secondary classrooms prior to the course. Based on concepts and ideas they learned in the CORP course, many participants believe they are better prepared to incorporate it in all courses for all their future students.

4.1.3.1 Participant perception of equity and proof writing: Prior to start of course

During the first interview, most of the participants (7 of 9) identified proof as a formal activity that was not taught by either them or their mentor teacher during their field placement. Eight of the nine participants said that they believe it is appropriate for all students. Three participants (Tanya, Karen, Lucy) explained that they enacted "reasoning like" activities since they asked students to explain why. Four participants (Tina, Uma, Katie, Katherine) explained that proof was not studied in their classrooms, but believed it should be include in secondary mathematics. An interesting finding is that even though most participants believed proof should be included in all secondary mathematics classrooms, several were less optimistic or specific about enacting proof tasks in their own classroom.

Only Brittany and Nathaniel spoke of experience with trying to enact proof tasks. Nathaniel was the only one to suggest that proof is more appropriate for honors level students. Brittany seemed conflicted saying that in general it is important, but through experience she wondered if it was appropriate for all students. Nathaniel talked about his experience and rationale with engaging students in writing proofs prior to the start of the course:

> *We actually did do some units on actually having the students make formal proofs, 2-column proofs. It's an honors geometry class, and so we try and bring proof in a little bit more because they are able to handle it and actually construct or actually articulate their arguments a little bit better than other students.*

Nathaniel's comment is forthright in explaining that he believes honors geometry students are better prepared to write and articulate proofs, and so he and his mentor teacher (the we) provided those students with the opportunity to construct proofs. He explained that he did not do proofs in high school and credits his experiences in college for his strong proof writing ability. However, he seems conflicted about what types of activities students should do in secondary classrooms when he says:

> *Simply engaging students in replication of procedures does not ensure that skills can translate to other areas; however, formal proofs may confuse students and actually hinder overall learning.*

Nathaniel was not the only participant to start the course with a formal view of proof, but was the only one to articulate that teaching formal proof construction might not have a place in the secondary education for all students.

Brittany believed that there should be more opportunities for students to write proofs in high school geometry. She said, "I wish there were more in geometry because there are so many theorems." However, when she spoke about her teaching experience with her mentor, Brittany explained:

> *Now as a teacher my students don't do two-column proofs in my geometry class. I teach geometry and college algebra. We don't really do any proving in algebra either. But we do - I mean we did at the very beginning of the year we did do the two-column proofs, but I found that my students weren't able to do it. Like they just – they struggle with it a lot no matter what kind of supports we were trying to give them. So I think we kind of backed away from that.*

Brittany did not directly mention proof being more for lower or higher tracked students, but explained that 'her students' were not able to construct proofs. In general, Brittany commented on the need for more proof in geometry, but the enactment of it was impractical with her students.

The majority of the participants (8 out of 9) were unopposed to the general inclusion of proof for all students but, as with Brittany, when the attention redirected toward their own classrooms they changed their belief in that all students are capable. Tina talked about her middle school students:

> *So like pre-Algebra, general math, I guess. I guess if I think about the students that I teach, there would be a barrier there just because of like reading comprehension and fluency, being able to put together a statement that clearly proves something. Being able to put together a sentence that gets across their thinking and how they're connecting to ideas. But I guess that's definitely something that you want in general in math is the ability to be able to connect to ideas and show why you can connect them. I guess that's kind of the whole point of how we've been taught to teach. Yeah. Definitely. It seems necessary in every kind of math if I think about it like that.*

The conflict is the tension between recognizing that proof is an important mathematical process for students to learn, and the reality that the students they teach find it difficult to construct proofs, so maybe it is not appropriate for 'my students.' Tina seems to become aware of the dichotomy as she is speaking and retracts her words to say that her students too should have the opportunity. Additionally, while the prospective teachers attempt to enact proof tasks in their mentor teacher's classroom, they recognize their limited pedagogical skills as well as with those of their mentor. This leads to the mentor teacher making a decision that they should "back away" from requiring students to produce proofs.

Other participants (3 of 9) shared how they value informal reasoning and explanation, but not did not see proof as part of the curricula or that they did not implement formal proof tasks. Lucy and Tanya shared similar thoughts while Karen addressed the curricula and grade level she taught. Tanya and Karen's explanations are shared:

> *Right now I don't do like formal proofs in my classes, but I do like a lot of justifying your work, so like I guess like right now I think that it takes a less like a less formal role but students should be able to explain and justify their work, but not necessarily write like a formal proof, write out a whole paragraph with each step of why. (Tanya)*

> *So as far as this year, I think since in seventh and eighth grade they don't straight out have to prove something but they have to reason through why two things might be*

equivalent or why something might be - you have to provide some justification even if that isn't a formal proof. (Karen)

The comments address the inclusion of justification in their classroom without specifying if it was more or less appropriate for certain ability levels. From these quotes, the take away is that prospective teachers believe making sense of mathematics is important for all students and this is attainable through informal communication of ideas. The quotes could also be interpreted to mean that they do not engage their students in proof because they do not believe it is appropriate or that their students are not capable. Tanya commented that she did not believe proof as she currently conceives it as a 'necessary' activity. Karen seems to be referencing her curriculum when she said that 7th and 8th graders do not have to prove mathematical statements. It is evident that some (3 of 9) participants believe informal sense making is an important part of teaching secondary mathematics, but students are not held accountable for constructing valid arguments.

Katherine, Katie, and Uma also spoke about the general need for writing proofs in high school without mentioning specific student ability levels. Katherine and Uma also did not share information about enacting proof tasks prior to the start of the course, so they both thought in general that students should have access, but it is not clear if they believe all students should engage in constructing proofs. Katie was the only one of the three that spoke personally about proof and her pre-service teaching experience in a secondary classroom. Katie said the following about her student teaching situation:

I really only finished my student teaching experience two weeks ago, so I really only had four months of being in the classroom, and it was a fairly traditional experience. The kids were not really asked to do anything out of the ordinary. Even when I asked them to explain things, they said "with words?" So it was very traditional, so no, we weren't able to really do too much with that. But if I have a chance in my own classroom, I think I would try to pull some of that in, because I think it's neat.

Katie's comment addresses the common struggle for many pre-service teacher placements where they are trying to promote explanation of ideas, but they are confronted with the established classroom norms set by the mentor teacher. Katie explains that justification was not a norm, so she struggled to encourage communication of mathematical ideas. However, Katie believes it is a practice she wants to establish when she is hired as a new teacher. Katherine, Uma, and Katie each conveyed their belief with the importance of having students engage with proof activities in secondary classrooms without mentioning ability levels, nor did they discuss situations with teaching students to reason or prove.

The design of the CORP material anticipated these teacher challenges. Researchers have reported that students struggle to write proofs even after successfully completing courses that require them to write proofs. Additionally, it has been reported that teachers are not prepared to support students with learning to write proofs. A course goal was to support the prospective teachers with knowledge of reasoning-and-proving and skills to enact tasks so that they are better prepared to support their students.

4.1.3.2 Participant perception of equity and proof writing: During and after the course

Since the participants' perception of proof as a formal product expanded to a set of activities that end with proof, most of the prospective teachers continued to recognize the utility of proof, but believed it is more accessible and applicable over time. Overall the participants recognized proof as an accessible addition to the informal explanation they were already encouraging in their classrooms, learned new ways to include proof, and believe they want to integrate reasoning-and-proving activities into their classrooms without specifying student ability groups. The final result is that the participants either believed that all students should have the opportunity to construct viable arguments or that no students should be provided a chance reason-and-prove, which in a sense is an equitable stance.

While all the participants recognized the range of activities associated with reasoning-and-proving as a way to prepare students to construct proofs, it is not clear that all participants came to believe that such tasks should be integrated into their curriculum based on what they said during the second and third interviews. Four participants were enthusiastic at the prospect with engaging all students' access to reasoning-and-proving tasks. Two participants make general statements about the importance about including reasoning-and-proving opportunities for all students and that it should be a priority, but reasoning-and-proving tasks do not seem to be something they personally will integrate into their future secondary courses. Three participants seem conflicted as to providing any students an opportunity to engage in reasoning-and-proving tasks. Also, eight of the nine participants did not identify proof as being an activity for a particular student ability group.

Four participants were enthusiastic with the prospect to integrate reasoning-and-proving tasks into their future classrooms. Tanya, Karen, Brittany, and Katie expressed an increased interest with implementing reasoning-and-proving activities in their classrooms. During the second interview Tanya said:

> *It's not something I've really thought about before this class. But now I think I wish we started learning this earlier because it would have been useful to do with my students throughout this year.*

So originally Tanya related her current practice of asking students to explain their thinking as similar to just an informal version of proof. After the first few classes Tanya comes to recognize the differences. During the third interview Tanya continued to say that she looks forward to incorporating reasoning-and-proving into her curricula as a new teacher.

Katie and Brittany both spoke about how they believed in the importance of proof and valued it prior to the start of the course, but now they have an even greater understanding of why it is important and how they would integrate it into the their curricula. They shared their thoughts during their third interviews and Katie's words below are representative of how both expressed their excitement about applying what they learned:

> *I think this class gave you tools of how you can implement into algebra classrooms, how you can better implement it in geometry, just by taking something that's even a more traditional question and sort of rewording it. Yeah I've always valued it, but I think this just gave you, you know a better way to do it.*

Prior to the course Brittany talked about how she and her mentor 'backed away' from having their students write proofs. Katie's quote highlights the point that it is not enough for teachers to want to or should integrate proof into their classrooms. Prospective teachers need to gain the knowledge to support their students successful engagement with proof. Katie and Brittany believe the course provided them with the 'tools' to implement what they believed was important.

Karen views it as her responsibility to prepare students for writing proofs and explained this during her third interview:

> *I think I just want to start from the beginning of the year, talking about how important it is to provide justification and to be thinking about how you can support your answers, so like to say, "Is that enough to convince a skeptic?" or whatever. I want to start with that right off the bat, saying things like that, to get students in the mindset of "How am I supporting what I'm saying? How do I know my answer's always going to work?" It's strange because that's something you have to start from at the very, very beginning. I really can't just start it in the middle of the year and expect everything to be perfect. It's definitely a process.*

Karen understands that developing students' ability to reason-and-prove will not be a simple exercise. As she explains it needs to be an 'ongoing process.' She is prepared and motivated to engage all of her future students to justify their mathematical thinking.

Two participants (Uma and Katherine) believe reasoning-and-and-proving is a beneficial activity for all students, but spoke in more general terms in that it is something mathematics teachers as a group need to integrate over the duration of courses without specifying if it is something they plan to do. Uma's quote below is representative of the general view that all teachers should include proof tasks in their classrooms:

> *Reasoning and proof is something that needs to be taught & students to develop over time. Students need to be aware of your expectations for what counts as proof. But in order for your expectations to mean anything to the student they have to develop an understanding of what counts as a proof what is needed what is sufficient. As a teacher, we need to scaffold their development of R&P skills by consistently incorporating it into the curriculum on a regular basis.*

While Uma and Katherine recognize what will be needed to prepare students, it is difficult to know if they view it as their responsibility to do so.

The final three participants (Nathaniel, Lucy, Tina) are conflicted about the possibility of including reasoning-and-proving activities in their future classrooms. For instance, even after discussing the purpose of proof as a useful skill that you want students to develop, Nathaniel still questions the benefit of having students in high school engage in the activity. Nathaniel identifies two concerns: 1) there is lots of content high school students need to study to build a foundational base, and 2) students might not be ready to develop proof arguments. Nathaniel shared the following thoughts during his third interview:

> *I would still say that I still have some questions about the benefits that the proof aspect could have to student learning. I've really come to see the reasoning because it gives students some understanding, and I can see how the proof will have some benefit, like holding students accountable and having them develop arguments, but I think sometimes*

> *the development of the proofs still might be too difficult that the foundation isn't already there for the students to think like that, or it hasn't been there for the past years, and it might be too tough*

He believes students need to learn lots of content or all content before they are ready to articulate proofs or he may be suggesting that high school students are not developmentally ready to construct proof arguments. Therefore, Nathaniel wonders if proof is appropriate for any high school student.

Tina and Lucy also identify covering content as a challenge with including proof. Tina is concerned that a future school district would not support the inclusion of reasoning-and-proving tasks. She explains that these tasks are worthwhile and practical, but feels it might be too time consuming and she also worries about keeping on pace with her peer teachers within the curricula. She shared her thoughts on this issue during the third interview:

> *I think it's more practical. But, you know, in like a district that didn't really see the need for it, it would be something that would be very hard to do. And although I don't think that you should necessarily try to go at it all by yourself you could – Even if it wasn't, like, the main focus of your curriculum you could sprinkle in tasks like these through the years and try to get students to think or see math in this particular way. That might help them actually like other things that are not necessarily taught where they have to prove something, but get them to think about things differently. …potentially more time consuming than other lessons that we've – other types of lessons that we thought about or planned throughout the year. But I would say just as worthwhile, if not more.*

Overall, Tina seems conflicted between student learning and identifying her role as a teacher, which is to follow a prescribed curriculum guideline. In addition to covering content, Lucy is not sure how to handle both formative and summative assessments with respect to proof. Lucy shared this quote on assessment, "How to grade student's proofs so I don't know. It seems like there is a very like fine line between what counts as a rationale and what counts as a generic example or whatever." Even though Lucy recognizes how the inclusion of reasoning-and-proving tasks can elevate student thinking beyond procedural skills what she calls ‘plug and chug,’ the challenge to assess along with the pressure to align with a mathematics department pacing guide is a real concern.

The point of this section was to uncover each participant’s conception regarding the level of students who should have access to writing proofs. Prior to the course, only Nathaniel verbalize the belief that honors geometry students are more qualified to engage in writing and articulating arguments. The remaining eight participants did not distinguish among ability groups either before or after the course. Of the eight who believed students are capable, four conveyed their increased understanding with how to support students and are interested to expand their knowledge with more teaching experience. The other four are not opposed to the idea. Two shared reservations about time and staying on pace with their peer teachers. The final two understand the commitment with preparing students, but spoke about incorporating reasoning-and-proving as something important for teachers in general, but not specifically identifying it as something they plan to do.

4.1.4 Opportunities: When (how often) should students engage in reasoning-and-proving activities?

This section focuses on which courses should include proof and how often students should write proofs in each of those courses. This section is related to the previous one in that if a teacher believes all students should be provided opportunities to construct viable arguments then it is important to learn if they believe that this means all students in all secondary courses and to what extent a course curricula should include opportunities for students to construct arguments. The point of this section is to learn if the course influenced the participants’ conception about proof being an isolated topic in a single high school course. The next section will report on the course the participants believe should include proof activities and the section that follows will explain how often throughout a secondary course the participants believe students should engaging in reasoning-and-proving activities.

4.1.4.1 What secondary courses should include reasoning-and-proving tasks?

The participants suggested that proof should be included beyond geometry even prior to the start of the course (as shown in table 4.3). Even though they started out believing proof could be included in courses beyond geometry, their focus changed from suggesting how it might be possible in all courses to discussing about how they would specifically integrate it based on what they learned in the course. For example, Nathaniel said that teaching proof is possible in all secondary content and at the end of the course explained that selecting and modifying tasks provided him a skill in which to provide students opportunities to construct arguments.

Prior to the course, only three of the eight participants mentioned that they thought proof was mostly for geometry students, but they extended the possibility to courses that follow geometry. Three participants could not imagine the type of problems students could prove in an algebra class. The other five participants suggested that they believed proof could be taught in all secondary courses. Seven of the eight participants changed their belief about the number of courses that could include proof based on their expanded view of what is proof and experience modifying tasks to include reasoning-and-proving.

Table 4.3. Participants' beliefs about which secondary courses should include proof

	In which courses should proof be taught?	
	Prior to the course	**After the course**
Nathaniel	Possible in all including elementary; not sure about benefits in any	Modifying and creating tasks showed me how reasoning and proving can be a very useful tool in developing students' ability to explain what is going on (still questioning the benefits)
Tanya	All high school courses (Algebra – Calculus) could include it	I think I will teach reasoning and proof from the very beginning of the year, regardless of the subject I teach.
Karen	Think everything specifically mentions pre-alg, algebra and geometry	Important to start from the beginning of the year and continue as an ongoing process
Tina	Probably most of them; algebra, essential to geometry, calculus even general math and pre-alg	I think that reasoning and proving is extremely worthwhile and should be attempted in secondary classrooms.
Lucy	Definitely geometry, not sure what they would look like in algebra and probably calculus	Even with a 'crappy' curriculum you can change questions without reinventing something to get at proof
Uma	Definitely geometry, not sure about algebra, probably Calculus	Two-column geometry are not the only kind; There's pattern-type tasks or any type of like algebra problem you could set up as a word problem
Brittany	Geometry and Calculus, not so much in algebra (basic stuff); Geo and above	I think reasoning & proof should be incorporated into every math course a student takes in high school.
Katie	All courses should include it (algebra through calculus even middle school	Learned tools to implement proof in all classes Algebra – Calculus
Katherine	(no record)	Reasoning and proving has long been relegated to geometry classes, but as the foundation of mathematical thought it should be taught at all levels.

At the end of the course, six of the nine participants specifically communicated that they believe reasoning-and-proving should be taught in all secondary courses. One participant, Uma, initially said definitely geometry, but after the course explains that she has come to understand how it can also be included in algebra. The three other participants did not specify courses, but similar to Uma explained how the reasoning-and-proving course taught them ways to provide students with opportunities to reason-and-prove. In fact, all of the participants explained how the course provided them with practical ways to modify existing tasks to include opportunities for students to develop arguments. For example, Tanya indicated that she wanted to include reasoning-and-proving in her future classes, but also during her third interview explained how modifying tasks will enable her to do it:

> *Okay, so it's [the course] influenced that by making me realize that it's more important than I previously thought to include reasoning and proving in my classroom, and like how I could do it, I've gotten better ideas about how I can do that, like with modifying tasks.*

Tanya hints at the fact that she may have wanted to include reasoning-and-proving, but know believes she has the practical knowledge to do it. Nathaniel and Lucy specifically address modifying tasks as an influence for how they can implement reasoning-and-proving in their classes as noted in table 4.3. Karen explains task modification as the 'big thing" she learned in the class, and others attribute it as how they will incorporate proof in any of the courses they may be assigned to teach.

Most of the teachers prior to the course suggested they proof could be included in all high school classes, and this view persisted throughout the course. However, they now believe they are better prepared with how to do it. Katie summarized this point in her third interview when she says:

> *Like I said before I mean I've always thought that it was important and not just at the college or calculus level. I think this class gave you tools of how you can implement into algebra classrooms, how you can better implement it in geometry, just by taking something that's even a more traditional question and sort of rewording it. Yeah I've always valued it, but I think this just gave you, you know a better way to do it.*

The participants believed teaching students how to construct and communicate arguments was important before they started the course, but now they have a better understanding with how they can choose and or modify tasks to implement with their students.

4.1.4.2 How often will the participants include reasoning-and-proving activities?

Since the participants are prospective teachers, their only experience teaching is with their mentor teacher during the past year in their field placements. So only a few of them spoke about trying to include proof as teachers in their current placements. Most of the participants said they did not teach proof at all, but did press their students to explain their thinking. Those they attempted to have their students write proofs commented on how it was an isolated topic. For example, Nathaniel explains how the textbook he used treated proof as an isolated topic, and the proof activities seemed to clash with the other exercises in the book:

> *I mean, cause I have limited experience even trying it, where our geometry curriculum already had it putting it in there, and even then it didn't seem to really match with the rest of the curriculum.*

Brittany was the only other prospective teacher to have experience teaching proof in a geometry course. She explained that she and her mentor teacher 'backed away' from proof lessons since the supports they were providing students were unsuccessful. Therefore, prior to the course the only participants who had tried to teach proof was in geometry classrooms and those were not situations were proof was integrated throughout the curriculum.

As was discussed earlier, at the end of the reasoning-and-proving course the participants felt better prepared to choose or modify tasks so that they can engage their students in constructing arguments. The rest of this section will report on how often they thought they would engage students in such activity in a secondary mathematics course.

Most of the participants expressed interest with integrating reasoning-and-proving tasks across their curricula (as shown in table 4.4). The categories are listed across the first row of the table show the options for integrating reasoning-and-proving from everyday to no integration of proof tasks. The column labeled 'isolated topic' is in line with the typical handing of proof in conventional geometry textbooks in which specific chapters are dedicated to writing proofs. The grouping of the participants is similar to how they were discussed in the previous section on equity.

Table 4.4. Participants thinking about how often they plan to implement proof tasks

	All concepts everyday	Integrate it throughout curriculum	Limited integration depending on time	Isolated topic	No integration
Nathaniel					X
Tanya		X			
Karen		X			
Tina				X	
Lucy			X		
Uma		X*			
Brittany		X			
Katie		X			
Katherine		X*			

None of the participants talked about engaging students in reasoning-and-proving tasks everyday. Those that eagerly discussed the importance of incorporating proof fell into the second category. However, two of the participants (represented by the asterisks) spoke about the general inclusion of proof into secondary classroom rather than speaking specifically about doing it themselves, which was discussed previously. For example, Katherine explained the placement of proof in secondary courses in the following way:

Students need scaffolding to understand what constitutes a proof, but should afterwards see that it can be integrated into any topics. Though extra work is required of the teacher, the development and modification of tasks will lead to greater student (and perhaps teacher) understanding.

Katherine and Uma both spoke similarly that 'the teacher' can support students with learning how proof is a process students' can learn across many content topics. However, it is unclear if Katherine or Uma plan to integrate reasoning-and-proving tasks across 'any topic' themselves.

The other four participants in the second category had a different tone about how 'they' planned to include it. For instance, during the third interview Brittany emphatically explained:

Okay. I definitely think whatever class that I'm gonna be teaching I'm gonna try to now incorporate it into the curriculum even if it has to be kind of like an extension off something but I think that we definitely need to make sure that it's in every math course in high school and I know it will be the first interview. I believed that in the first interview but I strongly believe it now.

Here Brittany is using the pronoun 'I', which provides a different level of responsibility from saying "a teacher" should. Additionally, she explains that she might use it as an extension to her curriculum as opposed to saying that it will be included in every lesson or that proof will be designated for a particular week or month. The only addition that Katie, Karen, and Tanya suggested beyond what Brittany stated is that they plan to start at the very beginning of the year.

The remaining three participants (Nathaniel, Lucy, Tina) are more skeptical of the possibility of including proof with their future students. Lucy falls into the third category (limited integration depending on time), since she identifies barriers to including proof so it seems as though it will be less of a priority for her. Lucy believes she is prepared to modify tasks and identifies benefits, but seems conflicted with the possibility. Her first quote seems to mean she is interested and aligned with the second category (integrate it throughout curriculum), when she says:

> *I said this before also that it's something that needs to be ongoing so it can't just be one lesson in the year. It has to be pretty consistent throughout the year for students to develop those skills.*

Then during the same interview her perspective seems to change as she talks about curriculum coverage as she explains:

> *Maybe how – I know this year for me pacing was a big issue. Or our curriculum was very – we had to teach a lot and we had very little time to teach it so maybe how you can incorporate these types of lessons and how you can get your students to think and reason like this when you are on a pretty strict time constraint.*

So Lucy understands the level of commitment needed to develop students' ability to reason and prove throughout a school year, but it seems as though this might become a secondary goal where the first is to cover the curriculum.

Tina also seems concerned about covering curriculum, but she says something specific that might mean she believes proof is a topic of study:

> *I guess when to do it, how early to do it, how often to do it, you know, how many of my tasks should I be changing to be more reasoning and proof like, how many – I don't think that every one of them should be like that. I think that'd be a little bit overwhelming. And I think they're fun, but definitely time consuming. So, you know, how much time do you spend pushing this reasoning and proving idea before you need to get back to doing something else or maybe a little bit less fun or interesting? I don't know.*

During the third interview, she begins to question and wonder about how often she might include reasoning-and-proving tasks and specifies that she would not include it everyday. Then she says reasoning and proving is an idea and how do you spend on it before returning to other topics. This quote seemed to suggest that Tina believes proof is not a process that needs to be integrated throughout various mathematical content areas, but a topic in itself.

Nathaniel's thinking about the teaching reasoning-and-proving was the most conflicting, making it difficult to gain insight with regard to what he may do as a classroom teacher. It seems as though Nathaniel will not include it at all when he said, "I would still say that I still have some questions about the benefits that the proof aspect could have to student learning." Nathaniel continues to think about the possibility and says that he is able to see how the reasoning connects to the types of thinking he promoted as a pre-service teacher, but he seems to believe even at the end of the course that writing valid arguments will be too difficult for high school students. He elaborates on his previous comment:

> *I've really come to see the reasoning because it gives students some understanding, and I can see how the proof will have some benefit, like holding students accountable and having them develop arguments, but I think sometimes the development of the proofs still might be too difficult that the foundation isn't already there for the students to think like that, or it hasn't been there for the past years, and it might be too tough*

He seems as though he is considering the inclusion of reasoning-and-proving tasks as he realizes the benefits, but in the end he seems to believe it would not be appropriate for high school students.

One of the goals of the course was to develop pre-service and in-service teachers capacity to integrate reasoning-and-proving tasks with their students. Seven of the prospective teachers never implemented proof tasks prior to the course and the two that tried did not believe they did so successfully. After engaging in the course, all are confident that they can implement reasoning-and-proving tasks, but just under half of the participants (4 of 9) are eager to get try enacting proof tasks, a couple of them (2 of 9) seem to believe it is important but seem reluctant to assume responsibility, and the final three participants identified challenges that may persuade them from attempting to implement proof tasks into any course they may eventually teach.

4.1.5 Summary of participants' conceptions of reasoning-and-proving

Interviews and in class written responses at the beginning, middle, and end of the course were analyzed to report on the participants' conception of reasoning-and-proving. Four predetermined categories were designed to capture the participants' perceptions of proof. The four individual conceptions will be summarized to portray a complete conception. The rationale for these four conceptions of proof is that if teachers have a full knowledge of what counts as proof, then they know the criteria in which they should hold students accountable. However, it is important to know for what reason students should construct proofs. So it is essential to broaden teachers' conception of proof beyond the narrow focus of a deductive organization of definitions and statements to include inductive reasoning to explore mathematical content. As teachers begin to expand their view of what counts as proof and how students can gain access, they may begin to recognize that it is appropriate for more students in more courses and that the benefits are such that students should engage in reasoning-and-proving more often.

All nine participants changed their conception of proof from a formal structure to include a variety of forms, representations, and types. While only one participant articulated all seven characteristics, eight of the nine participants articulated the three essential criteria of proof from the list. In addition, all of the participants expanded their conception of the criteria of proof by at least two of the seven characteristics throughout the course based on what they articulated prior to the course. The participants overall changed conception of proof seemed to impact their thinking about the purpose for enacting proof tasks as well.

Each of the nine participants expanded upon why they believe students should have access to constructing proofs. As a group the participants identified three purposes of proof prior to the course, and identified four additional reasons for engaging students in reasoning-and-proving. Researchers identified seven purposes for engaging students in writing proofs (i.e. Bell, 1976; de Villers, 1990; Hanna, 2000). The participants discussed several of the researcher purposes, but added two foundational reasons for secondary students: to learn what counts as proof and to support students in constructing argument that meet the expectation in the criteria. Since the participants developed an appropriate criterion for proof throughout the course, they now recognize the importance of choosing reasoning-and-proving activities to not only create a shared meaning for their classrooms, but to also choose tasks to scaffold students skill with constructing proofs. So if the participants believe they have the knowledge to choose tasks to provide students access, then the next two questions are which students in which classes should teachers engage students and how often (i.e. everyday, once a week, once a month etc.) during a school year should teachers choose reasoning-and-proving activities.

Research suggests that most teachers believe proof is an activity that should be relegated to honors students or courses with high ability students (Knuth, 2002b). Prior to the course, most of the participants (8 of 9) did not discuss how certain students are more capable or that proof should be reserved for a particular level of students. Of the eight who believed students are capable, four conveyed an increased understanding and personal interest with providing reasoning-and-proving opportunities. Two other participants understand the commitment with preparing students, but spoke about incorporating as something important for teachers in general, but not specifically identifying it as something they plan to do. These six participants who are generally interested or eager to include proof tasks, believe it needs to be integrated across all mathematical concepts. In addition, two shared reservations about time and staying on pace with their peer teachers. One of which may include it across multiple topics, but the other may view proof as an isolated topic of study.

The participants, prior to starting the course, defined proof, as a formal product that they thought could be included in courses outside geometry. The course expanded their conception of proof as a practice that can be included in courses outside geometry, and confident in their ability to modify tasks from whatever curriculum they teach to include reasoning-and-proving opportunities for all students. Therefore, proof evolved from a formal object toward a set of activities called reasoning-and-proving that result in constructing valid arguments that should be integrated into all secondary courses regardless of the student ability grouping level for most participants.

4.2 PRE-SERVICE TEACHERS ABILITY TO CONSTRUCT PROOFS

The results presented in this section correspond to the second research question:

> *2. To what extent do pre-service teachers construct valid and convincing arguments when prompted to write proofs over the duration of a course focused on reasoning-and-proving?*

The purpose of this research question was not only to identify which participants could write a valid argument, but also to examine the extent to which prospective teachers learned to reason-and-prove, including their ability to evaluate their own work. This research makes a distinction between a proof and a valid argument. All valid arguments are not proofs. A valid argument includes both proofs and rationales. A rationale is not a proof since it could include claims that require further explanation based on the community, or the mathematically correct argument fails to include statements to fully justify the conjecture. The participants' solutions to the eight tasks solved in class or during one of the three interviews were analyzed using the R&P codes, which were adapted from A. J. Stylianides and G. J. Stylianides (2009). This section analyzes the eight tasks while reporting on how the participants solved each problem. The participant challenges and demonstrated abilities are shared with respect to each task and summarized across the eight problems.

4.2.1 Analyzing participant results

The coding scheme described in chapter three was applied to all 71[6] solutions and the results are displayed in table 4.5. The names of the participants are listed in the first column and the eight tasks are listed in the order in which they were completed along the first row. The vertical shaded columns indicate the interview tasks. The codes A1 through A4 (A1: Example based or inability to make a generalization, A2: Incomplete or incorrect attempt to construct a general argument, A3: Valid argument but not a proof, A4: proof) represent the main argument category as described in the R&P codes. The number following the main code identifies the sub-code, which further specifies the type of argument. The trailing plus/minus symbols are applicable when the participant wrote a valid argument (i.e., they received either A3 or A4). The plus/minus symbols correspond individually from left to right to the three clear and convincing components: a) clarity in flow of argument including the use of symbols, language, and diagrams; b) clearly defined symbols and definitions; c) a clearly stated conclusion. A "plus" means the component was addressed in the proof, and a "minus" shows that the component was not addressed. The codes in table 4.5 indicate what each participant produced on each of the

[6] Karen was late to the class in which the Parallelogram construction task was solved and therefore did not do this task.

eight tasks. Some of the tasks were easier for the participants to write proofs than others as indicated in the last two rows of table 4.5.

Several themes emerge from reviewing table 4.5. There are no A0 codes and only three A1 codes among the 71 solutions. The absence of A0 suggests that the participants were, at a minimum, able to access the problem. Also the absence of the code A1.2 indicates that no participant presented a subset of examples and claimed they reached proof. Although some arguments were inductive, in most cases participants were able to produce a generalization. In each of the three instances where a generalization was not produced (code A1.1) the participant recognized the limitations of their solution.

Table 4.5. Results of participant solutions to all eight R&P tasks using the R&P codes

	1) N2+N is even	2) Squares	3) O + O = E	4) NxN square window	5) Parallelo-gram	6) Sticky Gum	7) Explain Number Patterns	8) Calling Plans
Nathaniel	A4+ - +	A4 - + +	A4+++	A4+++	A4 - ++	A4+++	A4 - ++	A4+++
Tanya	A3.1++ -	A4 - + +	A4+ - +	A3.1 + + -	A3.2++ -	A4+++	A4++ -	A4+++
Karen	A4 + - -	A2.5	A2.4	A2.3	N/A	A2.3	A2.5	A2.3
Tina	A2.4	A1.1	A2.4	A1.1	A2.1	A2.3	A3.2 - - -	A4+++
Lucy	A2.4	A2.5	A4+ - +	A2.3	A3.1+++	A4++ -	A3.2- - -	A2.2
Uma	A2.2	A3.1 - - -	A4+++	A3.1 + - -	A2.1	A2.3	A3.2++ -	A2.2
Brittany	A2.4	A2.3	A4+ - +	A2.3	A2.2	A2.3	A3.2- - -	A2.3
Katie	A2.4	A2.3	A4+++	A1.1	A2.2	A2.3	A2.5	A4+++
Katherine	A2.2	A2.2	A4+++	A3.1- - +	A2.2	A3.1++ -	A3.2- - -	A4+++
# of valid arguments	3	3	7	4	3	4	7	5
# of proofs	2	2	7	1	1	3	2	5

The majority of the codes (32/71) were of the A2 variety. In general, the A2 codes signify the participants' limited ability to explain why a generalization is always true. In some situations, the participants were aware of the limitations of their argument, but they were not able to improve it. However, in other cases, participants claimed to have produced a proof. Across the eight tasks, all seven participants that produced an A2 coded argument said that at least one was a proof. However, only two participants believed that each of their A2 coded arguments were proof. The lack of ability to evaluate the argument they produced demonstrates an inability to successfully apply their criteria of proof, which could lead to the prospective teachers accepting invalid student arguments as proof, which is what Bieda (2010) reported.

Combining A3 (13/71) and A4 (23/71) codes shows that the participants produced more valid arguments (36/71) than invalid (A1 or A2) (35/71). The A3.1 sub-code represents rationales and the A3.2 is applied when a conjecture was not stated and the participant assumed the statement they were attempting to prove. When creating the class criteria for proof, the participants agreed that all assumptions needed to be explained as something already proven. If the claim had not previously been proven, then it required further justification. However, recognizing assumptions within their own work was a challenge. In other words, all of the situations where a solution was coded A3, the participant believed she constructed a proof. Only two tasks (task 3: odd + odd is even and task 8: calling plans) did not yield any solutions that included assumptions. On the other hand, five participants produced solutions that included assumptions for the Explain Number Patterns problem.

While every participant constructed at least one proof, only two participants (Nathaniel and Tanya) were successful with producing more than two. A strong ability to develop valid arguments is defined as above a 75% success rate (7 or 8 valid arguments out of 8). Three prospective teachers (Lucy, Uma, and Katherine) demonstrated a moderate skill. Here, moderate is defined as those able to construct valid arguments on at least half the tasks (4, 5 or 6 valid arguments). The remaining four conveyed a limited ability, producing three or less valid arguments, which means they were unable to produce a valid argument on at least half the tasks (0, 1, 2, or 3 valid arguments).

Some tasks were more difficult for participants than others. For instance, only one proof was written for the Parallelogram Construction task, but seven out of nine participants wrote a proof for the "O + O = E" task. The lack of growth in participants' ability to write proofs may be due in part to the fact that tasks sampled different content knowledge (i.e. performance does not improve as more tasks are completed there are not more A3 and A4's later). It may be the case that the tasks were too different from one another for the participants to show improvement.

Finally, each of the 37 valid arguments was coded with the three clear and convincing "plus/minus" codes. Proofs to particular tasks were less likely to include minuses than others. For instance, many minuses are present for the Explain Number Pattern task, but none are listed in the Calling Plan arguments. No proof-code (A4) was followed by three minus symbols, but one non-proof valid argument (Lucy's parallelogram solution) was clear and convincing along all three measures. Some participants struggled with one of the three particular clear and convincing constructs. For instance, two participants (Uma and Tanya) did not include a concluding statement while writing at least three different arguments. As with constructing arguments, there is no evidence that participants improved along the clear and convincing dimension.

The following two sections: (1) provide an analysis of the eight tasks including the differences among them and the challenges that emerged across the participants' solutions; (2) summarize the challenges and provide possible reasons why growth was not detected.

4.2.2 Reasoning-and-proving task analysis

As each task is analyzed, representative solutions are shared to highlight successful solutions along with the challenges participants encountered, which will then be summarized across the eight tasks. Additionally, when known, a participant's evaluation of their solution will be provided. Finally, the tasks are compared to help explain why the participants may not have improved their ability to construct a proof from task one through eight.

4.2.2.1 N^2+N is always even

The first task participants were asked to prove is shown in figure 4.1. It was administered during the first interview and participants were not provided feedback on their work or asked questions to improve their argument. The task supplies a variable (n), defines it as a counting number, and the conjecture to be proven is provided. It is a typical problem students may have encountered in a college number theory course, in which exposure to high school algebra is the prior knowledge required to access the task.

Task 1: N^2+N is even: *Interview 1*
Prove that for every counting number n (1, 2, 3, 4 …), the expression $n^2 + n$ will always be even.[7]

Figure 4.1. N^2+N is always even task.

[7] Problem adopted from Morris (2002)

Overall, four different solution methods were employed to solve the task. Karen (A4 +- -) applied mathematical induction. Four participants (Tanya: A3++-, Tina: A2.4, Uma: A2.2, Brittany: A2.4) factored the expression into n(n+1), noticed that n and n+1 were consecutive counting numbers, and created two cases where n is either even or odd. Three participants (Nathaniel: A4+ - +, Lucy: A2.4, Katie: A2.4) employed a third method where they did not factor the expression, but defined a new variable to explain both the even (let n = 2k) and odd (let n = 2k -1) cases. A fourth method (Katherine: A2.2) was unique, in that a participant defined even as a number divisible by 2 (n = m/2), and squared the alternative form of the counting number before adding it to itself ($(m/2)^2 + m/2$). Finally, she factored out a two to show the expression is even. The four solution paths resulted in three valid and six invalid arguments, but the solution method alone did not determine validity.

Two of the valid arguments are proofs (Nathaniel and Karen) and one is not (Tanya), since the argument included assumptions. Tanya does recognize the use of the assumption, but does not further justify her claim. However, the most interesting comment of the three who wrote a valid argument came from Karen who used mathematical induction. The interviewer asked, "Why does the method of mathematical induction prove the conjecture? Karen admitted that it was a procedure she learned in college, but was unsure why it worked. All three claimed their argument proved the conjecture, but Karen was not sure why the method she used was valid other than her college professors telling her it was a viable procedure for proving.

The other six participants also claimed they proved the conjecture, even though they did not. Two of the six non-proof arguments were coded A2.2. Mathematically they did not define odd and even in a useful way to show the expression is always true. To highlight the challenge to define the terms even and odd, we can examine Katherine's solution:

$\Rightarrow$ *even*
$m = 2n \qquad n = m/2$

$(m/2)^2 + (m/2) = m^2/4 + m/2 = ((m^2 + 2m)/4)$

$= 2\ ((m^2 + 2m)/8) \Rightarrow n^2 + n$ *is even*

Katherine understood that her conclusion needed to show two times some quantity. However, she defined a counting number as any even number divisible by two. This definition did not provide her with anything different from what she was given. So instead of proving why the expression always worked, she manipulated the variable to produce an expression times two, which raises the question: Why is $(m^2 + 2m)/ 8$ a counting number? Therefore, the A2.2 codes represent solutions that were mathematically incorrect or unproductive with proving the conjecture.

The second noticeable challenge other participants encountered is related. They defined the terms even and odd in a way that could be useful in proving the statement, but were unaware that their algebraic argument did not account for all counting numbers (A2.4) or that they were defining variables several different ways in the same problem. Brittany's response is representative of how several participants struggled with using variables correctly to cover all cases:

Even number is 2n
$n^2 + n$
$n(n + 1)$

n = *odd*
$2n + 1$
odd • even number = even
↑
odd • 2n =
$(2n + 1)(2n + 2) = 4n^2 + 4n + 2n + 2 = \underline{2\ (2n^2 + 3n + 1)}$
↑
even number

n = even
even (odd)
2n • odd
2n(2n + 1)
$2\,\underline{(2n^2 + n)}$
even

Brittany writes that n is odd, but then on the next line writes 2n + 1. If she is defining n as all odd numbers, then only odd numbers are valid numbers to substitute into the expression 2n + 1. Mathematically this means that n is odd (1, 3, 5…), then 2n + 1 is a subset of the odd numbers (3, 7, 11…). Therefore, she is not proving the conjecture for all odd numbers, which conflicts with the way 'n' was defined in the problem statement.

Even if the issue of labeling 'n' as an odd number is overlooked, the problem with covering all cases is not resolved. Assume 2n + 1 represents any odd number, since n is given to be any counting number. This leads to the product Brittany wrote (2n + 1)(2n + 2). Substituting the smallest counting number in the product yields 3(4), which does not cover the case of the first counting number (one times two is not covered). Hence the argument is invalid.

Brittany and Katherine's solutions to the first task are representative of six of the nine participant solutions. They manipulated algebraic symbols and claimed the mathematical statement would always be true. While the given conjecture is true, their arguments were not clear about defining terms or variables. When an interviewer asked a participant why her argument was a proof, her response was because it was algebra. In other words, the prospective teachers (6 of 9) appear to believe that proof consisted of correctly manipulating algebraic symbols without attending to what the variables represent with respect to the problem situation.

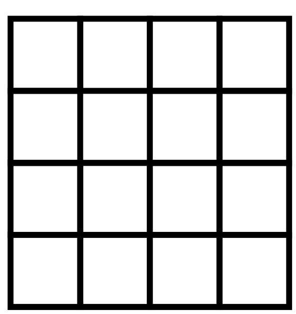
4.2.2.2 Squares Problem

The Squares task shown in figure 4.2 asks for the number of 3x3 size squares that can fit into a 60x60 size square. It is the only task among the eight that does not ask for a proof of a general situation. The first two questions in the problem encourage inductive reasoning since it requires the solver to first find the number of 3x3 squares in a 4x4 and then a 5x5 square before moving the larger case (question three) that would be cumbersome to draw and count.

The main difference between the Squares problem and the previous task is that the solver first needs to find a solution before explaining why it always works. In the $N^2 + N$ task it is given that the sum is always even and the participants were expected to either explain why it is true or find a counterexample. In general, in solving this task the participants followed two different inductive solution paths, which included: 1) using the smaller cases to make sense of how the 3x3 could move about the 60x60 square which leads to the correct solution (Nathaniel: A4 -++, Tanya: A4 -++, Uma: A3.1- - -, Lucy: A2.5) and 2) using the smaller cases to generalize the situation, using the general formula to find the answer for the specific (60x60), and then explaining why the answer must be true for the 60x60 case (Karen: A2.5, Tina: A2.4, Brittany: A2.3, Katherine: A2.2). These two solution methods are shared below with specific examples.

Task 2: The Squares Problem*: *Class 1

1. How many different 3-by-3 squares are there in the 4-by-4 square below?



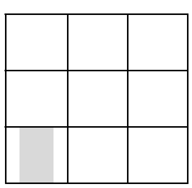
2. How many different 3-by-3 squares are there in a 5-by-5 square?
3. How many different 3-by-3 squares are there in a 60-by-60 square? Are you **sure** that your answer is correct? Why?

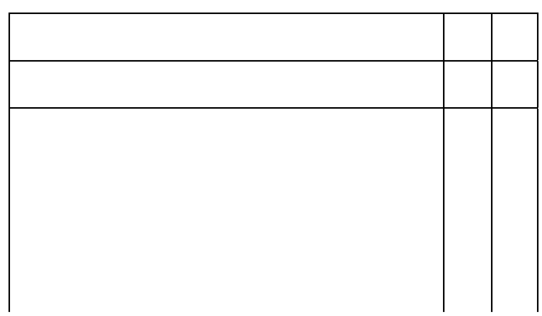
Figure 4.2. The Squares Problem

Nathaniel's solution is shared to highlight the first solution method described above, which leads to a proof:

> *All 3x3 squares have a bottom left corner. This square must have two squares to the right and two above it.*
> *The bottom left corner may not occupy the top two rows or the top two columns. In a 60 by 60 square there are 120 squares in the top two rows and 120 squares in the right most columns subtracting the 4 that are in both we get that there are 120+120-4=236 squares that cannot contain the bottom left corner.*
> *That's how (60x60) -236 = 3600-236 =3364 remaining. Hence there are 3364 total.*

Additionally, Nathaniel included two drawings in his solution, but neither was explicitly referenced in his argument.





Nathaniel uses the information in the problem to explain his methods for proving the problem situation. He first explains how he will use the 3x3 square to count within the 60x60 square. The drawings support the reader to understand why the top two rows and last two columns are not counted as he explains in his written argument even though he does not explicitly reference his pictures. He does make a labeling mistake with initially writing "top two columns," but corrects it in the next sentence writing "right most columns." This misstep along with not clearly explaining why he multiplied 60x60 is why the code reads A4 - + + for his argument. Overall the reader is convinced as to why 3364 is the correct solution since he used clear language and defined his terms to explain how he counted the total possible number of unique 3x3 squares that can be placed into a 60x60 square.

On the other hand, Karen is representative of those following the second solution path. She first drew out the first two cases 4x4 and 5x5. She drew nine 5x5 squares and showed all uniquely placed 3x3 squares within each 5x5. Then she created a table of values without labels as shown below:

3x3	1
4x4	4
5x5	9
6x6	16
nxn	~~$(n/2)^2$~~ $(n-2)^2$

60x60 ~~$(60/2)^2 = 900$~~

$$(n-2)^2 = (60-2)^2 = 58^2 = 3364$$

So instead of finding a pattern for how she counted the movement of the 3x3 square about the smaller six squares (4x4, 5x5, etc.), Karen looked for a numerical pattern in the table to make a generalization from empirical cases. She then used the generalization to find the correct answer to the problem. However, this method was not useful in justifying why 3364 is correct, since she was unable to justify how the length of the side the 60x60 square minus two [$(60-2)^2$] is connected to the problem situation.

The squares problem is part of a sequence of tasks where the goal is to learn that making a generalization from several cases is not a secure method for proving. In other words, the point of this problem is to teach participants, like Karen, that abstracting numbers from a few cases to justify a general case is not a viable method to prove this or similar situations. Therefore, it was expected that learners (practicing or prospective teachers and secondary students) would follow this method with the hope of learning through follow-up activities that this method is not secure.

Additional issues raised in this task include participants exhibiting a limited ability to articulate their thinking. Nathaniel explicitly chose the bottom left corner of the 3x3 square to support the reader with knowing how he counted the 60x60 squares. Others struggled to define terms in the problem and did not develop a complete argument. For instance, Uma writes, "you can only put the beginning of the 3x3 square in columns 1-58." There is no explanation about what exactly 'the beginning of a square' means. Brittany and Katie wrote about the number of shifts without specifying what was shifting or from where a shift starts. This problem required the solver to define terms to support the reader with understanding how they counted the movement of the 3x3 square, but many participants introduced new terms without clearly defining them.

4.2.2.3 Odd plus Odd is Even

More participants (7 out of 9) wrote a proof for the odd plus odd is always even task (as shown in figure 4.3) than any other task.

Task 3: O + O = E: *Class 2*
Prove that when you add any 2 odd numbers, your answer is always even.

Figure 4.3. Odd + Odd is Even Task

This task is similar to the first task ($N^2 + N$ is always even) in that the conjecture is given so the solver does not need to first find a solution as is with the Squares problem. However, this problem does not suggest the use of a variable. While most of the participants wrote a proof for this task, defining the terms odd and even along with appropriately choosing variables to show why the sum of any two odd numbers is always even was challenging for some participants. The question elicited multiple solutions. Most participants combined words and symbols as they did in their solution to the $N^2 + N$ task and most drew a diagram as a second method when encouraged to it another way. Only the first method was coded. The participants were asked to solve the problem a second way so that they could experience multiple solution paths and representations.

While seven participants constructed a proof, three were less clear with how they defined terms and or variables (Tanya: A4+ - +, Lucy: A4+ - +, Brittany: A4+ - +). The other four proofs were clear and convincing (Nathaniel: A4+++, Uma: A4+++, Katie: A4+++, Katherine: A4+++). Katie's clear and convincing proof for this task is as follows:

> *Odd numbers can be written in the form 2n + 1 because by definition they are not divisible by 2.*
> *Let $2n_1 + 1$ be one odd number and*
> *$2n_2 + 1$ be another odd number n_1, n_2 are integers*
> *then $2n_1 +1 +2n_2 + 1 = 2n_1 +2n_2 + 2 = 2(n_1 +n_2 + 1)$*
> *Thus, the result is divisible by 2 and is by definition an even number.*

Katie defined both n_1 and n_2 as integers and explicitly defined odd and even numbers. The argument is clear including a conclusion to justify the conjecture. A few participants defined an odd number as 2k + 1 and either did not explain what subset of numbers k represented or defined it as a natural number. Even though odd and even numbers are defined as integers, constraining an even number to the set of natural numbers is acceptable. In other words, focusing on the set of counting numbers and showing that the sum of any two odd counting numbers is even was accepted as proof for this problem.

Tina and Karen (both A2.4) are the only participants that did not write an acceptable proof. Karen defined the two odd numbers as n + 1 and n + 3. She defined n as an even number. Karen proved that the sum of any two consecutive odd numbers is even, which falls short of covering all cases. Tina defined any two odd numbers as n + 1 and n + 1, and went on to show that the sum is divisible by two. This shows the specific case of adding the same two odd numbers is even. The misunderstanding could be that a variable (n) can represent any number, so n + 1 and n + 1 are two different numbers. In other words, n can be any even number so one could substitute a six for the first n and 18 for the second n. Therefore, the odd plus odd problem like the $n^2 + n$ showed that some participants struggled to construct a proof since they exhibited a limited utility with defining variables.

More participants wrote a proof for this task than any other task. Some participants defined odd numbers as natural opposed to integers and a couple participants incorrectly defined any odd numbers as either the same number or consecutive odd numbers. This challenge to define terms and variables was more evident when participants followed a diagram solution method.

4.2.2.4 N-by-N Window Problem

The N-by-N Window problem as shown in figure 4.4 was administered during the second interview. The task first requires the solver to find the total length of wood strips for any size window prior to justifying why the generalization is always true. While similar to the Squares problem in that the participant needs to find and justify a solution, the NxN window problem requires a formula for any size window instead of a specific larger case. Five of the nine participants (Nathaniel: A4 +++, Tanya: A3.1 ++-, Lucy: A2.3, Uma: A3.1+- -, Katherine: A3.1-

- +) applied the method used to count the wood strips for smaller cases to find a generalization. Of the seven participants that generalized the N-by-N situation, three struggled to produce a valid argument (Karen: 2.3, Lucy: 2.3, Brittany: 2.3). Two participants (Katie: A1.1, Tina: A1.1) were unable to reach a generalization.

Task 4: N-by-N window: *Interview 2*

The diagram below shows the frame for a window that is 3 feet by 3 feet. The window is made of wood strips that separate the glass panes. Each glass pane is a square that is 1 foot wide and 1 foot tall. Upon counting, you will notice that it takes 24 feet of wood strip to build a frame for a window 3 feet by 3 feet.

1. Determine the total length of wood strip for any size square window.
2. Prove that your generalization works for any size square window.

3ft – by – 3ft Window

Figure 4.4. N-by-N Window Problem

Participants needed to seek a secure method for counting wood strips in order to provide a valid argument. Four levels of sophistication emerged in the analysis. The lowest level was to make tables of numbers (Tina: A1.1, Karen: A2.3). An argument needs to be based on how the windows are growing and a table of numbers is too far removed from the context to do this. Brittany related her generalization to the context, but does not provide an argument for why her generalization works for all cases. Four others (Tanya: A3.1++-, Lucy: A2.3, Uma: A3.1+- -, Katherine: A3.1- - +) made an attempt to provide an argument with varying levels of success. Katie (A1.1) tried to explain a secure counting method, but never reached a generalization.

Nathaniel (A4+++) reached the top level of sophistication since he constructed a clear and convincing proof.

Karen made a table of values and then encountered difficultly in making progress. Karen's solution is shared below:

Dimensions	*# of wood window strips*	
1 x 1	4	+8
2 x 2	12	
3 x 3	24	+12
4 x 4	40	+16

```
   n      n     n
3 + 4 + 3 + 4 + 3 + 4 + 3 = 24
n-1
```

4n + <u>n(n – 1) + (n – 1)n</u>
outside   inside

She realized the perimeter of any window was four times the length of the window. However, it is not clear why she labeled 4 with an n since n is three in that case. Basically, she found a correct generalization from the numerical values in the table, but was not able to justify why the generalization would always work for any size window. Karen admitted during the interview that her solution was not a proof.

The next level of sophistication was to concentrate on developing an argument for why the generalization always works from a specific case, in other words, construct a generic argument. Brittany never made a table of values and was clearer about explaining her generalization, but never developed an argument. Her solution is shared below:

3-by-3 window

O V H
4(3) + 2(3) + 2(3)

4-by-4 window

O V H
4(4) + 3(4) + 3(4)
N-by-N window
O V H
4n + n(n -1) + n(n – 1)
⇑ ⇑
# of wood strips will always be 1 less the # of panes
For any size window a square always has 4 equal sides

Brittany found a secure method to count the wood strips for two cases and then applied it to the general case, but failed to justify why the formula will always work. She labels individual parts of her formula without explaining why her method for counting will always work. Furthermore, Brittany believed this solution was a proof. In other words, she believed at this time that labeling parts of a formula that was generated from two cases was justification that it is always true.

Uma moves one step closer to constructing a proof since she did produce an argument, which is shared below.

nxn
4(n) *perimeter*
(n-1)n columns
(n-1)n

4n + n(n -1) + n(n – 1)
4(3) + 3(2) + 3(2)
12 + 6 + 6 = 24

For any size square window, the perimeter of the window = 4n, since a square is a quad w/ 4 equal sides.
When you divide the window into panes you create n columns by adding strips of wood. To create n columns you need n-1 vertical strips, & the strips need to be n ft long.
When you divide the panes into n rows you must also add strips to create n rows you need n-1 horizontal strips & the strips again need to be n ft long. Then add pieces together →
4n +n(n-1) + n(n-1)

Uma applies the definition of a square to explain why the perimeter for any size window is 4n, but does not define the variable n. She also does not justify why any size window will have n-1 vertical and n-1 horizontal strips of wood. Uma believed her argument was a proof even though she was unable to justify why there would always be n-1 rows and columns.

Nathaniel was the only participant to clearly articulate why his generalization will work for any size window. Even though his generalization is different, it is his explanation for why it always works that elevates his argument to proof. Nathaniel wrote the following solution:

> $n^2 + n + n^2 + n$
>
> *Assume nxn window. Then there are n rows of n panes. So there are nxn strips on the bottom of all panes and we add n for the strip on top. This gives us nxn + nx1 horizontal panes =*n^2+n.
>
> *The argument is the same for vertical panes. Thus counting the left side of each pane and noticing this counts all vertical strips but the right most side of the window frame.*
>
> *This will also give us* n^2+n *vertical wood strips for a total of* $2(n^2+n) = 2n(n+1)$

Nathaniel clearly explained how he counted the wood strips surrounding each pane without leaving the reader to wonder if his method would always work. The only vertical column of strips not counted is the right most one, which he adds to his argument. Specifically identifying how he counts the wood strips enables the reader to understand his thinking and why the counting works for any size window.

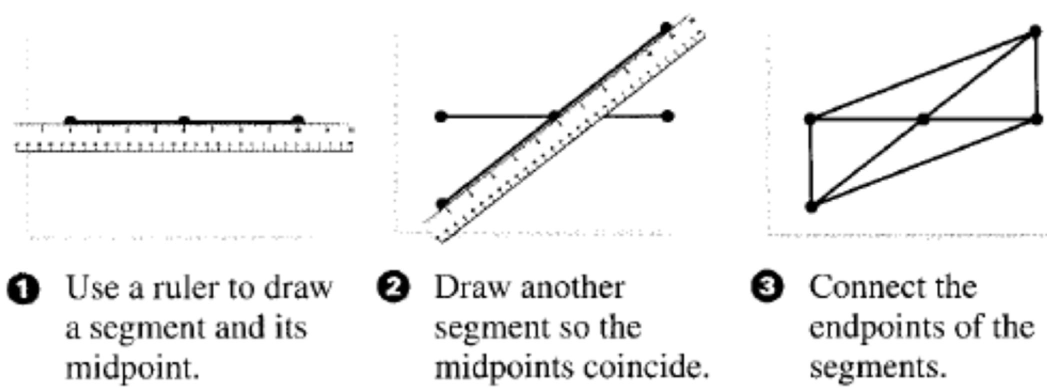
Starting with a table to find a formula, which Karen and Tina did, seems to interfere with promoting reasoning-and-proving. In other words, making tables of values supports students with a procedural or guessing method to reach a generalization, but the process of extracting numbers from the situation to place in a table does not foster a learner's ability to reason about the problem situation in general terms. The Squares task and the other problems in that sequence intended to foster participants' knowledge that it is not mathematically acceptable to derive general formulas to show that the generalization is always true from testing cases. It seems as though Brittany believes testing cases to derive a formula is an acceptable approach if the parts of the formula are connected to the problem context. The participants who constructed valid arguments relied on the diagram to explain how the total number of wood strips for any size window.

4.2.2.5 Parallelogram Construction Task

The Parallelogram construction problem is different from the previous four problems since it is a geometric situation and the task itself does not provide a conjecture. Based on a participant's construction, there is the potential to prove special parallelogram cases (i.e. rhombus, square, etc.). That is, there is more than one conjecture that could be made and proven. The task is shown in figure 4.5.

<u>Solve.</u>

Consider the construction below.



Use the construction with a variety of starting segments. What type of figure does the construction produce?

Using the results, make a mathematical argument that explains why that figure is produced each time by the construction. Be sure to provide reasons for your statements using axioms, properties, or theorems where appropriate.

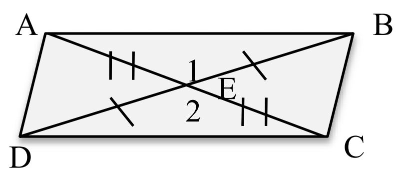
Create a new construction that also begins with a segment and its midpoint but is different in some way. What generalization can you make about any figure created by this construction?

(Adapted from McDougal Littell (2004), Geometry, p. 343, #29)

Figure 4.5. Parallelogram Construction Task

Question ‘a’ prompts the solver to identify the figure constructed and question ‘b’ requires the solver to create an argument that explains why this occurs. This implicit call for a conjecture was ignored by many (6 out of 8) of the participants. The task intended individuals to construct a variety of specific types of parallelograms based on the chosen construction. For instance, if the two line segments are bisected at a right angle, then a rhombus is formed. In other words, the construction becomes the conjecture for which to form an argument for why a particular figure is always formed. Most (6 out of 8) participants did not state how they constructed their figure. For example, Lucy wrote, “Conjecture: the construction always yields a parallelogram.” For the reader it is difficult to know what Lucy constructed or that Lucy knew that a parallelogram is constructed whenever two line segments intersect at their midpoints regardless of the angle formed and the length of the line segments. Other participants did not list a conjecture, nor did they provide a description of how their figure was constructed. For example, Nathaniel did not explicitly list a conjecture, but within his argument he explains that the intersection of the line segments create vertical angles and lists that each half of the line segments are congruent based on the construction of the figure. His conclusion explains that the construction is indeed a parallelogram. Tanya never wrote a conjecture and did not explain in the conclusion what she proved.

Even if it is assumed that all the participants understood the general constraints to produce a parallelogram, most (5 out 8) of them were not able to organize accurate geometric statements to prove the assumed conjecture. Uma's solution below provides an example of a typical argument that includes an error.

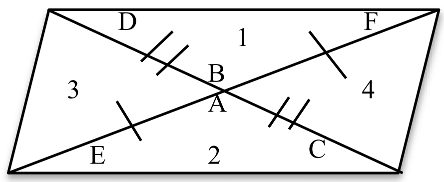


Produces a parallelogram

1. AE ≅ EC BE ≅ ED	1. Def of midpoint
2. ∠1 ≅ ∠2	2. vert. angle thm.
3 △AEB = △CED	3. SAS
4. AB ≅ DC ∠ABD ≅ ∠DCE	4. CPCTC
5. AB ∥ CD	5. Converse of ∥ line cut by transversal
6. ABCD parallelogram	6. def.

Uma's solution looks like a typical two-column proof. The statements are listed on the left and the reasons listed in the right column. As with most participants, she did not connect her construction with a conjecture to explain to the reader what she is proving. More problematic though is the error in the fourth line. It is true that the length of AB is equal to DC since the triangles are congruent, but angle ABD congruent to angle DCE is not a result of the triangles being congruent, nor do these two angles being congruent imply that AB is parallel to DC. This could have been a simple mistake where she meant to write BDC instead of DCE, but it could be a misunderstanding about congruent angles and which angles need to be congruent in order for the line segments to be parallel.

Where Uma's argument is an example of an A2.1 code, Katie wrote a solution that is representative of those coded A2.2 and is shared below.



Vertical angles ✏ A and B are =
SAS ✏ Δ ! Δ ✏ C and D are = ✏ E and F are =
Same follows for Δ 3 and Δ 4

Since < E = < F, the lines are parallel
SO we have a parallelogram

While Katie's argument is not organized in the traditional two-column format, it does provide reasons to support claims. However, the arrows she uses in the second line highlight illogical results. Katie's notation for triangles one and two should be congruent instead of similar, but the bigger issue is that she does not explain why the triangles are congruent. She wrote SAS, but it is not clear which sides and angles are equivalent. In other words, it is not logical to claim two triangles are congruent without explaining which parts of the two triangles are equivalent and how they came to be equivalent.

The Parallelogram construction task is unique in that it is the only geometry problem in the set and the participants were expected to state a conjecture based on their construction. Since geometry is where proof is typically addressed in secondary curricula, it might be reasonable to suspect that most prospective teachers would be familiar with the content and would demonstrate a strong ability to construct a proof for the task. However, the opposite was the case with these prospective teachers, since no participants connected their construction with a clearly written conjecture, and only one wrote a proof. The majority of them made mathematical errors or made claims without logically supporting how or why the claim is true.

4.2.2.6 Sticky Gum Problem

The Sticky Gum problem is most like the N-by-N window problem since it requires the solver to make a generalization and then justify why the formula works for any case. However, the Sticky Gum problem (as shown in figure 4.6) is more complicated since two variables are required in making a generalization. This task is also similar to the Squares problem in that it promotes inductive thinking for it requires the solver to first explore specific cases.

In solving the Sticky Gum problem, all participants were able to make a generalization of the situation, but many (5 out of 9) were unable to construct a valid argument. Most participants wrote convincing rationales for the first three questions, but many relied on explaining the parts of the formula as proof for the situation as was also noticed in the NxN window problem. It is important to note that the NxN window problem was an interview task, and the participants were not provided feedback on the accuracy of their work on interview tasks.

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 in 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.

Figure 4.6. The Sticky Gum Problem

Overall the participants followed two different solution paths (similar to those used on the NxN window problem). One method that four participants (Karen: A2.3, Tina: A2.3, Uma: A2.3, Brittany: A2.3) followed was to make tables of numerical values, which extended beyond the required cases outlined in questions one, two, and three to find a general formula. The second method included the participants' (Nathaniel: A4+++, Tanya: A4+++, Lucy: A4++-, Katie A2.3, Katherine: A3.1++-) reasoning from the first few cases and then extending their thinking to reach and justify the general case. None of the participants that followed the table method wrote a valid argument (similar to the NxN window problem). Katie was the only participant to follow the second method and not write a valid argument.

Tina's work is representative of the four participants who extracted values from the problem situation to build a table. Then she used the table of numbers to make a generalization, but does not explain why it always works. It seems as though Tina and others following this method first worked on finding a formula and then thought about how they might be able to create an argument. Tina's solution is shown below.

X	*Y*	
Kids	*Colors*	*Cost*
2	*2*	*3*
3	*2*	*5*
4	*2*	*7*
5	*2*	*9*

2n – 1 → # of kids

Kids	*Colors*	*Cost*
2	*3*	*4*
3	*3*	*7*
4	*3*	*11*

3n – 1
Let m = # of colors
n = # of kids

$$m\ (n - 1)\ +\ 1$$

⇓ ⇓ ⇓

# of colors | *one less than # of needed per kid* | *gumball to make the set*

The five other participants reasoned about the specific cases and used them to generalize the situation. It appears as though they thought about making a generalization and developing an argument as a connected activity versus first making a generalization and then thinking how might their formula be connected to the problem context. Then they explained how they envisioned the situation in general terms. Lucy's response below is representative of this second method to the Sticky Gum problem.

Let c = # of colors
Let n = # kids

If you have c colors, you could get each of the c colors in choices 1 → c. On the next choice you have to have one duplicate. To get another duplicate, however, you might have to choose

Assume there are n-many kids & c-many diff colors. We know that the maximum amount of money spent will occur if each color is drawn w/o any duplicates. Assume that each color is drawn without any duplicates. This will give you one of each color gumball. Assume that you again draw each color again w/o repeats. So then you will have two of each color. If this process continues and you choose each color n – 1 many times, you will have n – 1 many of each of the c many colors of gumballs. On the next choice (c(n - 1) + 1) you will get an nth duplicate of one color.

Instead of just thinking about how to find a general formula, she thought of the situation in a general way and explained her thinking. All of the participants that were successful with writing a valid argument followed a similar thinking process. In other words, the participants who made a generalization after constructing tables of values struggled to communicate the situation in general terms. They are able to explain the individual smaller cases, but did not reason about these smaller cases in productive ways in order to write about the situation from a general perspective. They did not show an ability to move beyond an inductive toward a deductive perspective.

4.2.2.7 Explaining Number Patterns Task

The Explaining Number Patterns task is similar to the Parallelogram construction task in that no conjecture was provided. However, unlike the Parallelogram problem, the Explaining Numbers Patterns problem explicitly requires a conjecture as shown in figure 4.7. The task also encourages inductive reasoning through example generation. The idea is that the solver would generate examples, notice a pattern, and then state a conjecture. The last question requires the solver to prove the conjecture.

Explaining Number Patterns

1. Pick any two consecutive whole numbers
2. Square each number and subtract the smaller square from the larger
3. Add the two original numbers together
4. Make a conjecture about the numbers you found in #2 and #3 (try more examples if you like!)
5. Prove that the conjecture you made in #4 will always be true. *After you have proven the conjecture in one way, see if you can prove it using another strategy or method.*

Figure 4.7. Explaining Number Patterns Task

The majority (7 out of 9) of the participants selected a variable without explaining how the variable related to the problem and did not write solutions to the first four questions. They created and manipulated equations similar to the $N^2 + N$ task. As was the case with the Parallelogram construction problem, the participants did not state a conjecture, so it was not clear what they set out to prove. Furthermore, a fully stated conclusion was also missing. Two participants did not complete their argument (coded A2.5), two others wrote a proof and the remaining majority ignored the call for a conjecture (coded A3.2).

Katherine's response (A3.2- - -), shared below, is representative of what the majority (5 out of 9) of the participants produced.

They are the same.
#2 $(n + 1)^2 - n^2 = n^2 + 2n + 1 - n^2 = 2n + 1$

#3 $(n + 1) + n = 2n + 1$

#2 = #3

Questions one through three did not receive a written response. The conjecture just reads that "they are the same" without explaining what exactly is the same. This is a concern since the conjecture could be written in a 'p implies q or q implies p' format. In other words, the converses are equivalent for this situation. The participants who produced similar solutions (Brittany, Uma, Lucy, Tina) also introduced a variable without defining it or explaining what it represents. While the task is unlike most of the problems they solved previously, the lack of explanation is surprising given that it was the last task completed in the course.

Similar to the "O + O = E" problem, this task asks for multiple solution methods, so most (7 out 9) participants moved from a method similar to Katherine's above to drawing diagrams. As was the case with previous diagram solutions, the variables were not clearly defined and the participants did not incorporate language to support the reader with understanding the diagram, variables, or the overall argument.

Finally, a few participants defined consecutive whole numbers inaccurately. When 'n' is the larger consecutive whole number and n – 1 is the smaller consecutive whole number there is a problem. Since a whole number includes zero (a lower bound), it is inappropriate to label the subsequent number as n – 1. For the case where n is zero, the number one less than zero is negative one, which is not a whole number. Therefore, again several participants demonstrated challenges with defining variables that represent abstract sets of numbers.

4.2.2.8 Calling Plans Task

The Calling Plans task was the final problem the participants were asked to solve (as shown in figure 4.8). It was most unique since it required a counterexample to prove the given conjecture. All of the participants were able to answer question one correctly, but ambiguity caused some participants to not write a proof for question two.

The Calling Plans Task

Long-distance Company A charges a base rate of $5 per month, plus 4 cents per minute that you are on the phone. Long-distance Company B charges a base rate of only $2 per month, but they charge you 10 cents per minute used.

Keith uses Company A and Rachel uses Company B. Last month, Keith and Rachel were discussing their phone bills and realized that their bills were for the same amount for the same number of minutes. Keith argued that there must be a mistake in one of the bills because they could never be the same. Rachel said that the phone bills could be the same.

A. Who do you think is right, Keith or Rachel? Why?

B. For any two phone plans, is there always a number of minutes that will yield the same cost for both plans? Provide an explanation to justify your position.

Figure 4.8. The Calling Plans Task

Focusing on question two of the Calling plans task, the majority (5 of 9) of the participants did identify a specific counterexample or a classification of counterexamples. For instance, Katie wrote the following solution.

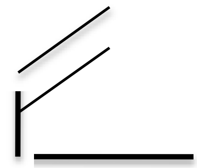
Parallel lines won't intersect – won't have a # of min. that yield same cost
Ex. y = 3 + .1x *x : # of min*
y = 5 + .1x *y: monthly cost*

Katie started with a classification of counterexamples, which is the case where the lines are parallel. She then provided a specific counterexample. This is all that was needed to answer the second question. However, some (4 out of 9) participants were either confused by the question or did not understand counterexamples.

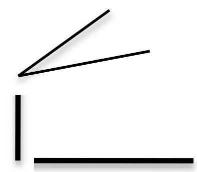
Those who did not provide a counterexample (Karen, Lucy, Uma, Brittany) answered the question similar to Brittany. They explained several possible cases for two calling plans without specifically answering the question. Brittany's response is provided below.

<u>Case 1</u>
Phone plans different rates / min. different monthly rates → 2 lines with different y – intercepts and different slopes → 1 intersection cost the same at some minute.

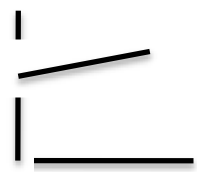
<u>Case 2</u>
Different monthly rates, same rate/minute
Two parallel lines and may will never intersect



<u>Case 3</u>
Same monthly rate and different rates/min



<u>Case 4</u>
Both parameters are the same
Always same



Brittany provided the four cases without explaining which cases or cases answer the question. Case one could be a solution if the two lines intersect beyond the maximum number of minutes in a month or intersect in the second quadrant. It is also not clear if Brittany believes that case three is a solution. If two people do not use a cell phone during the month they would only pay the monthly fee meaning their bills would be the same. Other participants provided similar solutions to Brittany, and it is not clear if they were confused by the question or the concept of a counterexample. Since this was the only task that assessed an understanding of counterexample it is not possible to compare the participants work on this problem with a previous problem.

4.2.3 Summary of participant challenges with reasoning-and-proving

The reasoning-and-proving tasks the participants solved surfaced the following mathematical challenges: 1) defining terms and variables, 2) transitioning from inductive to deductive reasoning, 3) stating conjectures, 4) clearly explaining thinking, and 5) indentifying a counterexample to prove a statement. The first four mathematical challenges were noticed across several tasks while the last one was only applicable to the Calling Plans task.

This research shows that prospective teachers encounter various challenges in writing proofs even if the content is at the high school level. None of the eight tasks required knowledge of mathematics beyond high school algebra or high school geometry, yet most were unable to construct proofs. In some situations the cause may have been a lack of knowing what is needed for an argument to count as proof, but some problems highlighted a lack of understanding about variables with respect to problem context. Therefore, it is clear that prospective teachers, even those with degrees in mathematics, need support with learning what makes an argument as proof and opportunities to construct proofs especially in secondary content so that they are prepared to supports students. The next five sections will discuss the challenges that surfaced across tasks.

4.2.3.1 Defining terms and variables

Some participants incorrectly or failed to define terms and variables on the following tasks: 1) $N^2 + N$ is always even, 2) Odd + odd is even, and 3) Explaining number patterns. The common thread across these three problems is that they are all number theory type problems. The context is abstract since it is situated in sets of numbers. Regardless if the problem providing a defined variable ($N^2 + N$ is always even) or the solver was expected to define variables, errors were present. If a set of numbers has a lower bound such as counting numbers, then it is essential to make sure how the terms are defined cover the entire set of numbers. In the case of the Explaining Numbers problem most participants did not even define the variable they choose. It was also evident that more support is needed when prospective teachers incorporate diagrams. The terms and variables were rarely defined in diagram solution methods.

4.2.3.2 Transitioning from inductive to deductive reasoning

Several participants struggled to transition from examining examples to providing a deductive argument. This issue was evident with the following problems: Squares, NxN window, and Sticky Gum. The common feature of these problems is the examination of specific cases to make a generalization, and explain why the generalization is always true. The issue always included making a table to generalize the situation. The participants who demonstrated this challenge were able to understand and explain the smaller cases that they could visualize followed by extracting the numbers from the smaller cases to make a table of values to develop a formula. The final step was to explain and connect parts of the generalization to the problem context. The problem with this method is that the generalization is based on a few cases meaning it lacks justification for why the formula will always work. The alternative is the explain why a specific case will always which was discussed in the course as a generic argument or extrapolate what is changing in the smaller cases to justify why they will always hold true.

4.2.3.3 Stating conjectures

Making a conjecture can take on several forms. When a conjecture is in the form of a formula, the participants knew to state the generalization. However, when the conjecture should have taken the form of a statement such as an if-then statement, the conjecture was not stated. The participants were asked to state a conjecture for both the Parallelogram construction task and the

Explaining number patterns problem. Of the eighteen solutions to these two problems only one conjecture was stated. Prospective teachers need more support realizing the importance with clearly stating what it is they are proving or it may be the issue that they need more support learning how to write out a conjecture.

4.2.3.4 Clearly explaining thinking

Clearly articulating thinking was also a concern. In certain situations it was difficult to discern between mathematical errors and simple writing errors. This was most relevant to the Parallelogram construction task. Additionally, participants introduced new nonmathematical terms through arguments without clearly explaining what the words mean. This was noticed while reading arguments to the Squares, NxN window, and Sticky Gum problems. If the issue is a simple writing error, it could be corrected simply by asking teachers to reread their work. The remaining errors are either mathematical or a limited ability to articulate thinking, which if addressed as an issue during instruction may improve.

4.2.3.5 Identifying a counterexample to prove a statement

Only the Calling Plans task required a counterexample to prove the statement and since the problem was solved during an interview, there was no opportunity to learn from it. Regardless of the fact that this problem was solved without a previous attention to counterexamples, it would be expected that secondary mathematics teachers should know what a counterexample implies and when it is appropriate to use one to prove a statement. However, a few participants appeared to not understand the role of a counterexample. It could be based on the wording of the question, so this concept requires more attention.

4.3 PRE-SERVICE TEACHERS ABILITY TO ANALYZE STUDENTS REASONING-AND-PROVING SOLUTIONS

The results in this section are in response to the third research question:

> 3. *To what extent do pre-service teachers improve their ability to distinguish between proof and non-proof arguments created by students over the duration of a course focused on reasoning-and-proving?*

There were a total of 32 student solutions across five tasks that the nine participants were asked to classify as proofs or non-proofs. During each of the three interviews after the participant solved a task they were asked to judge the validity of four or five selected student arguments to the same task ($N^2 + N$ is even (n=5), NxN window (n=4), and Calling Plans (n=5)). In addition, after solving the "O + O = E" task (n =10) and Sticky Gum problem (n =8) the participants were asked to analyze solutions. Here 'n' refers to the number students solutions analyzed for each task. The 32 total student solutions were selected for several reasons including the opportunity to question teachers' conception of their criteria of proof along with supporting teachers in gaining experience with anticipating the wide array of solution methods. Therefore, this section will share the results of the participants' ability to distinguish between the proof and non-proof arguments to all solutions, and then concentrate more specifically on the student solutions that were chosen to assess teachers' understanding of the criteria for proof.

4.3.1 General analysis of student solutions and participant results

The results for each of the nine participants are shown in table 4.6. The first number in each ratio indicates the total number of solutions that the participant correctly categorized as 'proof' or “non-proof”. The denominator represents the total number of solutions the participant critiqued. All ‘unsure’ responses and lost data did not contribute to the total. So each participant was asked to evaluate 32 solutions, some data was lost and some participants said that they were ‘unsure’ if the argument was or was not a proof. For example, the audio recorder stopped during Tanya’s first interview, which is why she only critiqued 31 arguments. Brittany was the only other participant for which the data was lost due to recording issues. Her last two responses in the third interview were not recorded, and Brittany also said that she was unsure twice while evaluating the odd + odd is even solutions. The two responses lost to collecting data plus the two ‘unsure’ replies combines to four non-responses, which is why Brittany is only credited with analyzing 28 solutions. The other participants with less than 32 analyzed solutions listed are a result of them saying ‘unsure’ about their decision with labeling the argument.

Table 4.6. Percentage correct while evaluating reasoning-and-proving solutions

Nathaniel	Tanya	Karen	Tina	Lucy	Uma	Brittany	Katie	Katherine
28/32 (88%)	27/31 (87%)	24/31 (77%)	25/29 (86%)	26/29 (90%)	27/32 (84%)	21/28 (75%)	29/32 (91%)	26/30 (87%)

All of the percentages were rounded to the nearest whole number. Each participant correctly identified three fourths or more of the arguments. Katie and Lucy each only misidentified three solutions while Karen and Brittany each incorrectly analyzed seven. While some of the student solutions were proofs and others non-proof arguments, the solutions varied with respect to the extent to which they addressed the criteria of proof.

To answer this third research question, the participant results will be shared for all 32 solutions. Then the types of solutions that were easy and more challenging for the participants to identify will be discussed. It is important to note that a "correct" analysis of a student solution is not just picking between proof and non-proof. The participants needed to provide an accurate explanation. Finally, to determine growth, a deeper analysis of the more challenging solutions is reported. The results fall into three categories: 1) expected to be easy and they were, 2) expected to cause a challenge and they did, and 3) expected to be a challenge but were not.

4.3.2 Analysis of the student solutions and participant results

The participant group results are displayed in table 4.7 for each of the 32 student solutions. The tasks are listed chronologically in the first column in the order in which participants encountered them and the letters across the first row are associated with the individual solutions for each task. For instance, the interview one task ($N^2 + N$ is even) included five pieces of student work for the participants to evaluate (A – E). The ratios in each cell signify the total number of participants who correctly identified the argument as proof or non-proof to the total number of participants who responded to the solution. The blank cells represent solutions in which all the participants correctly identified the tasks as proof or non-proof. The '♦' identifies the solutions which are considered distracters, which will be discussed in more detail later in this section. The P and NP signify whether the solution is a proof or non-proof respectively. For example, the cell for task 1 ($N^2 + N$ is even) student A includes the ratio 7/9 and a P. This code means that the solution is a proof (P) and that seven of the nine participants correctly identified it as a proof and therefore two participants identified it as a non-proof argument. The 'U' listed in some cells represents instances in which a participant was unsure of whether or not the solution was a proof. The number before a U indicates the number of participants that were unsure about the solution. Only one student wrote that they were unsure of a solution within the set of Sticky Gum student solution set (solution H), and all the other instances in which a participant used "unsure" referred to the "O + O = E" student work. All nine participants correctly evaluated 12 out of 32 solutions. Furthermore, five "O + O = E" solutions were correctly evaluated by eight of nine participants where the ninth person was unsure. Overall the participants demonstrated a strong ability with analyzing most of the solutions, and struggled as a group to correctly identify several 'distracter' (♦) solutions.

The participants collectively critiqued 288 student solutions while data was collected on 285 solutions. Overall 234 responses matched the designated labeling[8], making for an 82% (234/285) success rate. Furthermore, 11 of the responses were coded 'unsure', and 40 responses disagreed with the argument categorization meaning that only 14% (40/285) of the responses were incorrectly labeled.

---

[8] The CORP development team identified the 18 student solutions analyzed in the course, and the author validated the 14 students solutions in the interview protocols.

Table 4.7. Total participant responses to each of the 32 solutions

Student Work:	A	B	C	D	E	F	G	H	I	J
1) N^2+N is even	P (7/9)			NP:♦ (4/9)	P:♦ (7/8)	______	______	______	______	______
3) O + O = E	1U	P:♦ 2U(5/9)		1U	1U	1U	1U		NP:♦ 3U(5/9)	
4) NxN square window	NP:♦ (7/9)	P:♦ (8/9)	P (8/9)		______	______	______	______	______	______
6) Sticky Gum			NP:♦ (3/9)				P (8/9)	NP:♦ 1U(7/9)	______	______
8) Calling Plans	NP:♦ (4/9)	P:♦ (6/9)	P:♦ (6/9)		NP:♦ (1/8)	______	______	______	______	______

The 40 (40/285 or 14%) incorrect responses were spread across 15 student solutions in which seven are non-proofs and eight are proofs. The participants were more likely to label a non-proof argument as proof than the opposite case. The eight proofs resulted in 71 (8 solutions times 9 participants with one missing piece of data) recorded participant responses of which 13 did not match the researcher labeling for an 18% (13/71) disagreement. On the other hand, the seven non-proofs were misidentified 44% (27/62) of the time. Therefore, based on the 15 of the 32 solutions where participants incorrectly analyzed student solutions, they were more willing to label non-proof arguments as proof than the proofs as non-proofs.

While the analysis in this section paints a broad view of the participants' ability to identify student solutions with respect to proof, focusing on the actual solution and what the participants communicated provides a deeper understanding of the results. The next section will share the types of student solutions that the participants were successful at identifying along with those that were more challenging.

4.3.3 Participant understanding of applying criteria of proof

The general results reported in table 4.7 show that the participants as a group seem to be successful in applying the criteria for proof on many student solutions and struggled to do so on several others. The types of solutions were designed to span the spectrum of possible argument types (i.e. empirical, rationale, generic, demonstration). While all 32 solutions will not be analyzed, themes will be discussed followed by examples to support the claims.

4.3.3.1 Student solutions, which were straightforward to identify

Some secondary teachers are convinced by empirical arguments (Knuth 2002a), as are some secondary students (Healy & Hoyles, 2000). Therefore, student solutions were selected to learn if prospective secondary teachers were convinced by example-based arguments. The findings suggest that the participants were not convinced by examples.

The first set of student work ($N^2 + N$ is even) and the third set (Sticky Gum) included example only solutions and no participant labeled these as proofs. The $N^2 + N$ student work contained two solutions with example only solutions. The problem expected the solver to prove why any number when substituted into the expression would yield an even number. Morris (2002) found that prospective elementary teachers were convinced of the example-based solutions such as the one shown below in figure 4.9.

<u>Argument 2 – Student B (Ben's Solution):</u>
Let n = 1. Then $n^2 + n = 1^2 + 1 = 2$. 2 is even, so this works.
Let n = 2. Then $n^2 + n = 2^2 + 2 = 6$. 6 is even, so this works.
Let n = 3. Then $n^2 + n = 3^2 + 3 = 12$. 12 is even, so this works.
Let n = 101. Then $n^2 + n = 101^2 + 101 = 10{,}201 + 101 = 10{,}302$. 10,302 is even, so this works.
Let n = 3056. Then $n^2 + n = 3056^2 + 3056 = 9{,}339{,}136 + 3056$. 9,342,192 is even, so this works.

I randomly selected several different types of numbers. Some were high, and some were low. Some were even and some were odd. Some were prime and some were composite. Since I randomly selected and tested a variety of types of counting numbers, and it worked in every case, I know that it will work for all counting numbers. Therefore, $n^2 + n$ will always be even.

Figure 4.9. Student B's solution to $N^2 + N$ is always even

None of the participants in this study identified Ben's solution as proof. Additionally, they knew why it fell short of being classified as such. For example, during the interview, Lucy explained her rationale; "He proved it worked for a number of choices, but not for all. They don't prove that it always works." Lucy's criteria of proof included that an argument must apply to all numbers and recognized that what Ben produced was only convincing for the numbers he checked. Lucy's response is representative of the group of participants. While some students or prospective elementary teachers are convinced by examples, these participants were able to apply their criteria of proof that an argument must cover all cases and were able to recognize that student B only checked a few examples.

Empirical arguments are not limited to a set of examples and can extend to include a generalization. The Stick Gum problem included two solutions that based a generalization on examples, but the participants recognized the limitations of these arguments. The Sticky Gum student F solution (shown in figure 4.10), which is shared below is an example of a generalization.

F

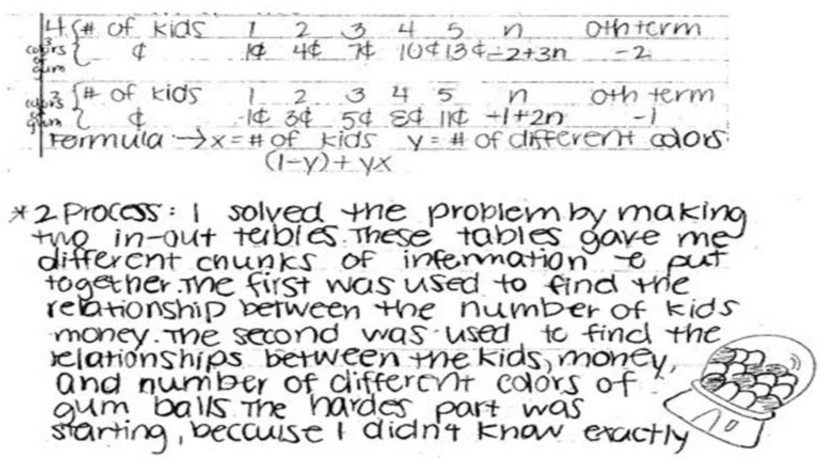


Figure 4.10. Sticky Gum student solution F

The student who produced solution F reached a correct generalization and explained their process. However, all nine participants identified the argument as a non-proof argument and were able to explain why it was not a proof. For example, Tanya reasoned: “Found a pattern in tables, but does not explain it in general.” Tanya’s rationale is similar to the others in that a generalization needs an explanation for why the formula will always work. The generalization is an extension of the example only based solution, and they are both considered empirical arguments. By the end of the course the participants knew that these types of empirical arguments are not proof.

The participants knew prior to the course that examples were insufficient for proof, but it is possible that some learned that generalizations are not proofs. A goal of the second task and first in course (Squares problem) was to support the participants with understanding that a generalization based on a set of examples is not proof. There was no assessment of a generalization as proof prior to the participants engaging in the Squares problem.

4.3.3.2 Student solutions, which were more challenging to identify

In order to assess teachers’ conception of proof, the student solutions were designed to include argument types that were expected to cause conflict among the participants such as: empirical arguments with justification, generic arguments, counterexamples, and rationales. Also different representations (i.e. words, diagrams, etc.) were mixed among the four argument types to challenge participants to consider form and representation in creating a proof.

A distracter is an argument that is likely to challenge a community of learners in coming to agreement on the classification as a proof or not a proof. The participants in this study recognized that example only based solutions or reaching a generalization were not proofs and there was no disagreement so those types of solutions are not distracters for this group. Based on the first iteration of the course materials and a review of research, twelve distracters were identified that included non-proof (rationale and empirical) and proof arguments (generic argument and counterexample). It is important to emphasize that this study distinguishes among various types of empirical arguments where adding justification to a generalization based on examples is more sophisticated than an empirical argument comprised of examples only. As shown in figure 4.11, it is the 'gray area' that includes argument types where disagreement occurs. Depending on the group such as elementary pre-service teachers or undergraduate mathematics majors, the 'gray area' may include different argument types.

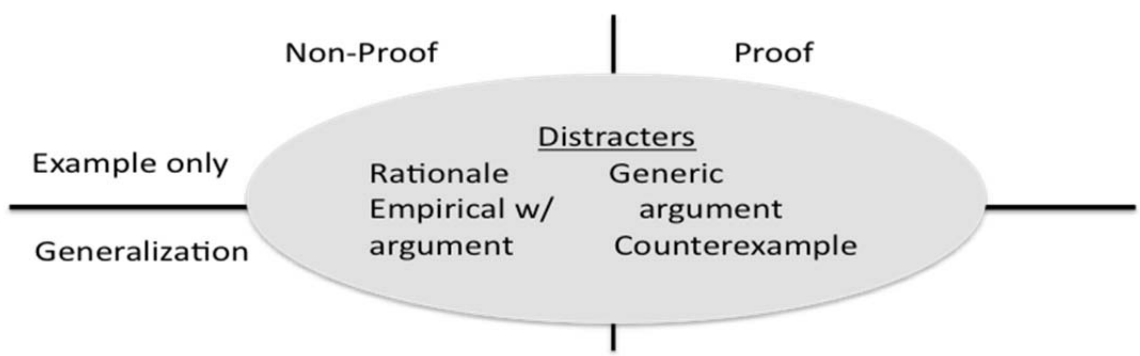


Figure 4.11. Types of arguments that were distracters

In addition to the argument type, the solutions were intended to press on teachers' conception of representation. A traditional view of proof would generally mean that a valid argument would not include narrative language or pictures. The distracters not only vary along argument type (i.e. rationale, empirical argument, generic argument), but also include pictures and narrative language. The twelve student solution distracters are listed in table 4.9 along with the ratio of correct to number of participants that responded and the argument type. All twelve tasks listed in table 4.9 correspond to the distracters '♦' labeled in table 4.8.

The task names are listed in the second row and the student solutions are listed vertically below each task name. For instance, the student D and student E solutions in the first column are solutions to the $N^2 + N$ is even task. The argument is listed below the student identifier. Below the explanation in each cell is the ratio of the number of participants that correctly identified the argument. So at least one participant disagreed with each of the twelve distracters. There were only three non-distracter labeled solutions in which participants disagreed.

Table 4.8. Twelve pre-determined distracters

Twelve Identified Distracter Student Solutions				
1) N2+N is even	**3) O + O = E**	**4) NxN square window**	**6) Sticky Gum**	**8) Calling Plans**
Student D: Rationale: The argument makes several assumptions. (4/9)	Student B: Generic argument: makes a general claim from a specific diagram example 2U(5/9)	Student A: Empirical: constructs a generalization from a few examples without justifying. (7/9)	Student C: Empirical: Provides justification for one case. (3/9)	Student A: Not valid An incorrect narrative response. (4/9) Student B: Counterexample Provides a narrative general and a specific counterexample. (6/9)
Student E: Generic argument: makes a general claim after examining specific diagram examples (7/8)	Student I: Rationale: correct statement, but assumes too much. 3U(5/9)	Student B: Generic argument: uses a particular case to generalize to any size window. (8/9)	Student H: Empirical: Provides justification, but not general to all cases. 1U(7/9)	Student C: Counterexample: Provides a convincing response while mentioning the counterexample. (7/9) *Student E:* Not valid: Provides a general counterexample argument with errors (1/8)

Rationale: There are two rationale distracters listed in table 4.9: student D ($N^2 + N$ is even) and student I (O + O = E). These solutions are mathematically correct and are intended to assess the participants' ability to identify assumptions. While the goal is not to reprove every statement in every argument, the message is for teachers to support their students with constructing proofs that draw on previously proven concepts (A.J. Stylianides & G.J. Stylianides, 2009).

More than half the participants incorrectly labeled the student D (shown in figure 4.12) solution during the first interview.

Student D (Dominique's Solution)

If n is an odd counting number, then n^2 will be odd. An odd plus an odd is even, so since n^2 and n are odd, $n^2 + n$ is even.
If n is an even counting number, then n^2 will be even. An even plus an even number is even, so since n^2 and n are even, $n^2 + n$ is even.
Since all counting numbers are either even or odd, I've taken care of all numbers. Therefore, I've proved that for every counting number n, the expression $n^2 + n$ is always even.

Figure 4.12. Student's D solution to $N^2 + N$ is always even

This argument is considered a rationale and not a proof based on the four imbedded assumptions: 1) *If n is an odd counting number, then n^2 will be odd,* 2) *An odd plus an odd is even,* 3) *If n is an even counting number, then n^2 will be even,* and 4) *An even plus an even number is even.* Each of these claims should be justified or explained as previously proven.

Five participants labeled this argument a proof. They focused on the mathematical correctness and how the argument was similar to or different from how they approached the problem without attending to the unsupported claims. For example, Karen's response is typical of the five participants that identified student D as proof:

So that would be kind of general –starting with the case that I did where it would have been the next step so that it happened for every time to generalize it. And she did the same thing for even. And so then she's saying since our choices are only even or odd, she's taken care of all of the different possibilities. So, yes, that's a proof.

It seems clear that Karen is applying her current criteria of proof which includes being general, mathematical correct and covering every case. However, she did not attend to the assumptions. The other four participants drew attention to the assumptions as to why the solution falls short of proof. For example Tina said:

There is nothing in Dominique's proof to show that an odd number plus an odd number is always going to give you an even number. Whether that's a definition of something or a theorem or a postulate or something, that would need to be included.

Tina is explicitly questioning one of the claims in the argument for justification on why the solution is a non-proof argument. Therefore, even though the argument is valid, it fails to support all the claims or does not explain that the claims were previously proven.

The second rationale solution is student I (O + O = E) (shown in figure 4.13), which also challenged the participants. While five correctly identified the solution as a non-proof, only Nathaniel labeled the one sentence solution as a proof. The other three participants said they were unsure.

Student I

If you add two odd numbers, the two ones left over from the two odd numbers (after circling them by twos) will group together to make an even number.

Adapted from: Coxford, A. F., Fey, J. T., Hirsch, C. R., Schoen, H. L., Burrill, G., Hart, E. W., et al. (2003). *Contemporary mathematics in context: A unified approach: Course 3*. New York, NY: Glencoe McGraw-Hill.

Figure 4.13. Student I solution to the "O + O = E" task

Even though Nathaniel wrote yes that student I is a proof, he was not overly supportive of the solution. He wrote "conditional" followed by "enough info (information) is present; however, it probably should be explained in better detail." Brittany, Lucy, and Katherine all wrote that they were unsure, and explained that the argument was correct but suggested that it did not fully convince them of the truth as to why an odd plus an odd was always true. For example, Katherine wrote, "The logic is sound, but the student is unclear about what she means by 'the 2 ones left over from the 2 odd numbers'." Therefore, all four participants that did not label the solution as a non-proof argument held considerable reservations about the strength of the argument.

Most of the participants did not seem concerned about the lack of algebraic symbols in either of these arguments. Brittany was the only participant to make a comment that could be related to the informal structure of the argument. While reviewing the student D solution, Brittany explained that she would like "more math." It is difficult to know if she needed more explanation or she wanted to see more symbols in place of the language. Comments about the lack of algebraic symbols did not resurface again while analyzing solutions after this task, but the inclusion of symbols will be discussed again later in this section.

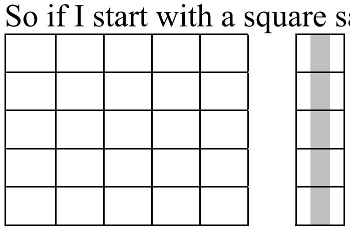
It is difficult to know if the participants improved their ability to identify unsupported claims. During the first interview, several participants' criteria of proof did not require the justification of assumptions. This requirement was specifically addressed after the participants analyzed the "O + O = E" student work, which included the student I solution. The class community came to an agreement that assumptions need to be thoroughly explained or labeled as previously proven, but there were no rationale arguments in subsequent student solution sets to test whether this had been learned.

Generic argument: A generic argument is a proof that justifies the general features of a mathematical situation while examining a specific case (G.J. Stylianides, 2008). There are three generic argument distracter solutions, all which include a picture. Only one participant disagreed that both Student E ($N^2 + N$ is even) and Student B (NxN window) solutions were proofs and two participants incorrectly identified Student B (O + O = E) as a non-proof. The reasons for disagreements varied, but none of the participants labeled the arguments non-proofs based on the inclusion of a picture.

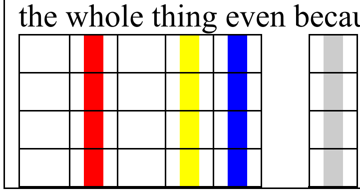
The first time the participants were exposed to a generic argument was during the first interview (task 1) when they were asked to analyze student E's solution (as shown in figure 4.14). The solution is generic since the odd case is based on the specific number five and the even case on six. This solution combines narrative language and diagrams, which was intended to question teachers view of what is an acceptable representation for proof. Seven of the participants accepted the generic argument as proof, data on one participant was lost, and the remaining student identified the solution as a non-proof.

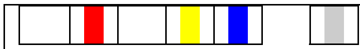
Student E (Edward's Solution)

So if I start with a square say 5 by 5 and add it to the number 5

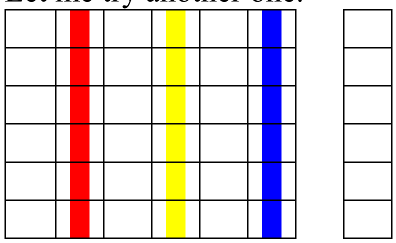


Ok now I will match up the colums so that all but one column has a pair (the blue one). The blue column will be matched with the gray 5 coulmn that is added to the square. So that will make the whole thing even because you can divide the entire thing into two equal pieces.





Let me try another one.



The columns in the 6 by 6 match up perfectly with none left over and the added part 6 folds in half. So every number is paired which makes $6^2 + 6$ an even number.
Now I got it. If the square is an odd by an odd like 5x5, then there will always be a column left over since an odd number does not divide by 2 evenly. The left over column of an odd sided square will always match with the added column part.
If the square is even by even, then every column has a match. The added part for an even by even will also be even based on the problem. And an even number divides two with nothing left over or folds perfectly.
So it does not matter the counting number that you start with when you square it and add it to itself it will always result in an even number.

Figure 4.14. Student E's solution to $N^2 + N$ is always even

The student E solution first examines two cases before moving to the more general case. The two cases are not explicitly labeled, but reach a conclusion where the odd and even cases are connected to justify the conjecture. The definition of even and odd is also embedded through out the language and diagrams. Therefore, the specific case of five is used to generalize the odd case and six is the special case to explain the odd, which is why student E's solution is a generic argument and proof.

Brittany was the only student to not accept the generic argument, but seemingly for a different reason. She explains that she likes the visual approach, but claims it needs to be more formal. Brittany explains her discomfort when says:

> *It's different from the other ones because they try to do it visually. I think the proofs have to be more formal. I think as an informal proof this might be good. But if we're talking about a formal proof, then I don't think this would be justified by just trying – again he tried some examples but – and he is tying to generalize for any case that would either be*

odd or even. I mean I like it but I wouldn't say this is the strongest proof that we could have done for this question.

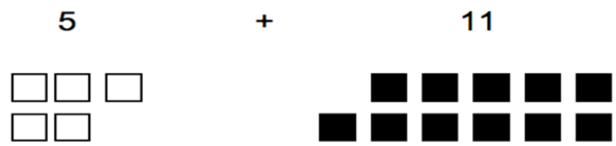
Brittany's analysis makes it difficult to know exactly why she is saying that it is not a proof, but she may be saying that a proof cannot include a picture. Immediately following the comment about visual she says that proofs need to be more formal. She does not comment on the lack of symbols or the use of language, but this could be interpreted to mean that formal proofs are not suppose to be visual. It is also possible that she is uncomfortable with the use of examples when she says that he tried examples that represent cases and follows this statement with the word 'but.' Overall it seems as though Brittany is uncomfortable with both features purposely designed to be distracters: 1) the use of examples and 2) the use of a diagram.

Even though the seven others said yes student E wrote a proof, some participants examined the two distracting features before making a decision. Some focused on the use of specific examples, but were convinced with the transition to generalizing the situation to all cases. None of the participants thought the diagrams were problematic. Uma explained that it was not needed, but thought it provide beneficial support with understanding the language. Tina was the most conflicted about accepting the generic argument, but finally decided that the picture persuaded her to accepting the argument as proof. She directly considers the definition of generic argument when she says, "I think he's very confusing because they're using a picture of a specific instance but talking about it generally." So even though the idea was not discussed prior to evaluating the student E argument, Tina was considering the definition of generic argument and wondering if it is acceptable as proof. After reading through the argument several times she said that she would accept it while commenting that the diagrams were necessary for her to understand the argument.

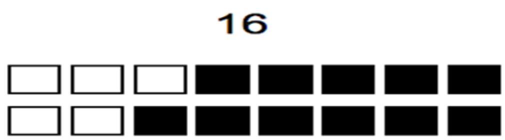
The student B solution (O + O = E: task 3) (as shown in figure 4.15) utilizes a specific example, but talks about it in general ways.

Student B

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., et al. (2003). *Contemporary mathematics in context: A unified approach: Course 3*. New York, NY: Glencoe McGraw-Hil

Figure 4.15. Student B solution from the "O + O = E" task

Brittany was the only participant to comment on the use of a specific example. She indicated that she was unsure and provided the following reason:

> *I can't say yes or no because the student uses a specific example to generalize, but then does not mention this is true for all different odd numbers. Maybe a less specific diagram would prove it.*

Clearly, Brittany is applying her knowledge that a proof must cover all cases and she is not concerned about the use of the picture. However, Brittany is not convinced about the use of a specific diagram. It may seem that she is contradicting herself when she says that the specific example is used to generalize, which is the definition of a generic argument, but the students were not introduced to the concept of generic argument prior to analyzing this student work ("O + O = E).

Three other participants were also either unsure or labeled the solution a non-proof. Tina, Tanya, and Lucy accepted the use of the specific example to generalize, but did not believe that defining odd as having 'an extra block' as acceptable. For example, Tanya wrote, "Does not prove that extra blocks always form a pair." In other words, the diagram of specific numbers five and eleven was accepted, but the fact that these two numbers have an "extra block" implies that all odd numbers will have an "extra block" was not accepted which is a valid criticism. A critical criterion of proof is that it needs to be accepted by the community, and based on how the term odd may have been defined (if at all), the classroom teacher needs to make a judgment what is and is not accepted.

While four participants did not label the argument as proof, it was only Brittany that disagreed or was conflicted with the idea of a generic argument counting as proof. The other three participants focused their disagreement on the way the student solution defined an odd number. Therefore, all of the participants accepted pictures, and only Brittany did not view the generic argument as proof prior to the concept being formally discussed.

The participants examined the third and final generic argument (task 6) after they learned about the reasoning-and-proving framework where they considered the various argument types. As was explained in the previous two solutions, most participants were not distracted by the use of specific example to generalize the situation and most were not concerned about the use of language or diagrams. This continued to be the case for the student B solution of the NxN window problem. However, there are two points worth noting: 1) Lucy labeled the solution as a non-proof, and 2) Brittany not only accepted the argument as proof she called it a generic argument.

Lucy labeled the student B argument a non-proof because she was not convinced with how the student generalized the argument. Lucy labeled the argument a rationale. She wanted more explanation to support the claim. Again this is a legitimate complaint for her to recognize a weakness in how the student articulated their thinking and she expected more clarity. In other words, Lucy found a point of contention in the argument and was not directly disputing the use of a specific example to generalize or that the argument failed to be considered a proof since the student used a picture.

On the first two examples (student D: task 1 & student B: task 3) of a generic argument, Brittany claimed the solutions were non-proof because of the use of pictures and or specific example. While analyzing the student B and student C solution (NxN window: task 4), Brittany explicitly categorized the solution a generic argument, which she said is a proof. The three generic argument student examples with different representations did not distract most of the participants. Brittany was the only one to state that a diagram or using examples was too informal causing the solution to not count as proof. Finally, most of the participants (8 of 9) correctly identified the Sticky Gum student G (generic argument) as a proof, and the one disagreement was not related to the use of a specific example or the inclusion of a diagram, it pertained to the clarity of the argument.

<u>Empirical Argument</u>: An empirical argument is an example-based argument that fails to generalize the situation for all cases. A generic argument is a proof that is generalized from a specific case. This can be confusing since they both include examples, but the distinction is that the argument needs to be general for it to count as proof. Two types of empirical arguments

were previously discussed (example only and generalization from examples) did not cause conflict among the participants. However, there exist additional types of empirical arguments not yet discussed that are different from one another, and from the types previously analyzed. The first is the student A (NxN window) solution, which is a formal argument that includes algebraic symbols in which it is expected that teachers would accept based on the use symbols. The second is student C (Sticky Gum), which is a less formal solution containing language and a diagram.

The student A (NxN window) (as shown in figure 4.16) solution is not a proof since it generalizes a situation based on five examples without explaining why the pattern will always work. Seven participants recognized the limitations of the argument. Tanya was able to connect this example to what she learned on the first day of class. She explains that the generalization is based on the several examples in the table and then makes the connection to a class activity:

> *So like it could be a different pattern like if he would have looked at the second problem that we looked at on the first day, and wrote out a table for the first five. He could have come up with a pattern that will miss, and it would be wrong, is like a similar thing could happen here so I don't think this is a proof.*

Tanya recalled the Circle and Spots problem, which was the second problem she solved during the first class. The pattern detected in examining the first five cases does not continue in the same way from the sixth case on. Other participants made similar comments with how the solution falls short of proof since it does not explain why the pattern will continue in the same way.

Student A

Window size x	# of wood pieces y
1 by 1	4
2 by 2	12
3 by 3	24
4 by 4	40
5 by 5	60

I notice that the first difference in the table is 8, 12, 16, and 20 and all of the second differences are 4. Since the second difference is constant (4), then the equation is quadratic.
I know that the y-intercept is 0 since a 0 by 0 window will have zero wood pieces.
Also half of the second difference gives the leading coefficient. Now I just need to find the coefficient for x, which I will call b.

$$Y = 2x^2 + bx$$

Choosing a random coordinate will allow me to find b. So I will choose (2, 12).

$$12 = 2(2)^2 + b(2)$$
$$12 = 8 + 2b$$
$$4 = 2b$$
$$b = 2$$

So for any square size window length x, the number of wood pieces is

$$2x^2 + 2x$$

Figure 4.16. Student A solution to NxN window

The algebraic procedures in the student A solution, however, did cause three participants to rethink their decision and two participants decided to identify the argument as a proof. Three participants (Tina, Karen, and Brittany) talked about the procedures in the problem as possibly being previously accepted in the community in which it was constructed. In other words, if the methods used in the solution are accepted classroom methods that it should be acceptable as proof. However, they failed to realize the generalization was based on five examples and the argument does not justify why the pattern will continue to be quadratic. Karen, after considerable thought, changed her mind as she explained:

> *I don't think so because they found it but they didn't explain why it worked, and so I don't think finding the equation is showing why it works and why it's always going to hold. So I don't think it is. I got thrown off because they were doing all these like math things.*

Karen realized she was distracted by the solution because of the algebraic procedures or what she called "math things." Then she reconsidered her criteria of proof, which included the need to explain 'why' something works in which this solution did not do. Therefore, the algebraic symbols were a distraction for a few participants and the symbols cause two of the participants to not recognize the insecure method used to generalize the situation even though it was directly attended to during the first class meeting.

The student C (Sticky Gum) (as shown in figure 4.17) solution is also an empirical argument. The solution reaches a generalization, but describes the generalization using a specific case.

Student C

Here is the formula needed to rewrite problem 4 algebraically:

x = colors

y = children

z = cents

xy – (x – 1) = z =

3 • 3 – (3 – 1) =

9 – 2 = 7¢

The reason I chose this formula is as follows. I needed to multiply the colors by the children in order to get the maximum amount of money needed (including children getting more than one color of the same color). But since the children only have to have the same color as one of the gum balls, I needed to take away the other two possibilities, which is why I subtracted the color minus 1. Look at the following diagram:

X:	1¢	1¢	1¢
Y:	1¢	1¢	1¢
Z:	1¢	~~1¢~~	~~1¢~~

See, we don't need the last two results, of the triplets getting the same color of all the gum balls, just one color – which is why we subtracted the last two numbers, by taking the number of colors, and subtracting one, which in this case is 3 – 1, giving us two, which we subtracted from the kids times the colors, resulting in 3 • 3 – (3 – 1) = 9 – 2 = 7. Whew! Long sentence!

Figure 4.17. Student C solution to the Sticky Gum problem

The shortcoming with the Student C solution is that the explanation of the generalization uses the specific three children three-gumball color case instead of using the specific to explain the general case. The second sentence is general where it reads, "multiply the colors by the children," but the very next sentence explains that they needed to subtract "two." Also the general statement does not justify why you would want to multiple colors by children or what the product would mean. The "long" sentence below the diagram does not make any attempt at being general. Therefore, it is not clear how the generalization is reached, but it is explained using a specific example (empirical argument) opposed to generalizing the situation from a specific case (generic example). This nuance caused conflict with deciding on how to identify the argument.

None of the participants were distracted by the use of the narrative language or pictures to make a decision. However, seven participants were unable to initially recognize that the language is situated in a specific case, or they have a misunderstanding of a generic argument. Only Katie and Nathaniel recognized the fact that this is a non-proof argument. Lucy first wrote yes (proof) then wrote no, but her reason suggests she originally thought it was, but the class discussion persuaded her to change her decision to non-proof. Katie wrote out her reason for why the solution is insufficient, "[The solution] focused on a single example – needs to extend to a general case to hold true for all possible cases." Katie and Nathaniel understood that explaining a generalization based on a single case does not count as proof and it seems as though Lucy agreed with their rationale.

Since the reasons are only a sentence or two long, it is difficult to discern differences among the six other responses. In other words it is challenging to know if the remaining participants believe they that student C's argument was an empirical argument or if they thought it was acceptable to use an example to explain a generalization. It seems as though Karen belongs to the former perspective when she writes, "Generalizes based on an example, explains the variables and explains the colors – 1 part." It appears that she views the solution as a generic argument. Brittany may belong to the other perspective since she wrote, "Shown for general case and have explained why each piece of the formula exists." Brittany may believe that reaching a generalization and explaining it is a proof. While both views are incorrect, believing the solution is proof is a greater concern since it reveals a general misunderstanding of proof where the other perspective is a misinterpretation of this specific solution.

These student solutions highlight the complexity with evaluating teacher's understanding of empirical arguments. While they all recognize that examples alone and a generalization with little to no explanation is not a proof, a few participants were distracted with the use of algebraic symbols. Determining the validity of an argument was most challenging for the participants when an argument includes generalization even if the explanation does not include all cases.

<u>Counterexample</u>*:* The term counterexample was introduced on the first day of class as part of the sequence of three tasks, but understanding what a counterexample means was not the focus of the set of activities. The learning outcome for the three task series was to understand

that a solution based on any number of examples is not a proof. The next time the concept of counterexample was introduced was during the final interview. The problem included a false statement and the expectation was to find an example to prove the statement false.

More than two participants incorrectly identified four out of five arguments. In this section, the analysis will be on the two most controversial solutions (Student A and E). Both of these solutions are identified as non-proof arguments, but the majority of participants labeled each solution a proof. The focus of this analysis is on part B of the calling plans since this is question in which a counterexample is needed to solve the problem.

Student A's solution (as shown in figure 4.18) is not a proof since it failed to include the monthly fee variable. If the monthly fee is the same in two plans, and the cost per minute is the same in both plans, then the plans are identical and will always cost the same.

Student A

A. I think that Rachel is right because both Company A and B cost $7 for 50 minutes. I figured this out by making a table.

#	Cost A	Cost B
0	5.00	2.00
10	5.40	3.00
20	5.80	4.00
30	6.20	5.00
40	6.60	6.00
50	7.00	7.00
60	7.40	8.00
70	7.80	9.00
80	8.20	10.00
90	8.60	11.00
100	9.00	12.00

B. Any two phone plans that don't have the same cost per minute will be lines that intersect. If they have the **same** cost per minute they will be parallel lines that never meet.

Figure 4.18. Student A solution to the Calling Plans task

Another issue that was not addressed by any of the participants is the error in the first sentence of the argument. The graphs of two linear functions with different slopes will intersect, but only an intersection in the first quadrant makes sense in this context. Negative time is not a realistic quantity. In other words, when plotted, two calling plans could intersect in the second quadrant and never share the same total cost for the same number of minutes.

Four participants labeled the argument as a non-proof, but only three of them provided legitimate rationales. Tanya commented that she would like more information without specifically commenting on what she thought was missing. The three other participants (Lucy, Tina, and Katie) recognized that the general case was problematic since student A did not account for the possibility of the monthly fee being the same.

Katherine and Nathaniel recognized the limitation, but accepted the argument as proof anyway. Katherine first recognizes the situation in which the same cost per minute and same monthly fee would mean the two plans are identical, but she further explained that her analysis might be too critical. At this point, Katherine explains that she believes student A is providing a counterexample 'of sorts' in recognizing that parallel lines would be a situation in which two plans would never cost the same. In summary, Katherine and Nathaniel recognized the limitations of what was written, but assumed the student was aware of the issue that contradicts their solution.

The remaining three participants (Uma, Brittany, and Karen) seem to have a limited knowledge of a counterexample. They focused on explaining all the possible situations and believe a valid solution must attend to multiple cases. They claim that since student A is only focused on the case in which the monthly fee are different and the cost per minute are equal and graphically this would produce two parallel lines. They labeled this argument as not a proof because it did not covering all of the other possible situations. So regardless of the fact that they correctly characterized the solution, their response incorrectly explains why the argument is not a counterexample.

It is challenging to know what Karen, Uma, and Brittany understand about a counterexample, because they each make conflicting comments. For example, Karen first labels student A as a proof, but then after analyzing the student C solution she returns to the student A solution to change her mind saying, “Yeah A isn’t a proof because of the same thing that I did, didn’t talk about this one [a situation where two plans would not have a common cost], this case as well.” It seems as though Karen learned that a proof must cover all cases and is applying that rule to a false statement, so she does not seem to understand that a false conjecture only requires a single specific instance that disputes the claim. Brittany, Karen, and Uma each changed their decisions several times while evaluating the Calling Plans solutions so it is difficult to know if it was the question they did not understand or what it means to prove a false claim. At one point Brittany did seem to recognize that the situation only required a specific example, but mentioned contradiction, counterexample, and generic argument in the same sentence as if these terms are related somehow. Therefore, the overall issue may be that they are conflicted about applying their criteria of proof to a situation where it does not apply.

Student E (as shown in figure 4.19) was labeled proof by seven of the eight[9] participants. As with the other Calling Plan solutions, the response focuses on a general argument opposed to providing a specific counterexample. The part B question for the Calling Plans task is as follows: “For any two phone plans, is there always a number of minutes that will yield the same cost for both plans?” The student E response focuses on when two plans do intersect while

[9] The audio recorder did not capture Brittany’s response to the student E argument.

writing "No" they will not always be the same. The answer only focuses on when the two plans share a common total cost for a specified minute during a month. While it does not seem as though student E believes that the general case shared is the only possible situation for the problem, he or she did not explain any situation in which any two plans would not yield the same cost.

Student E

A. Rachel is right. I used my graphing calculator and put in the two equations

$CA = .04m + 5$

$CB = .10m + 10$

And found that the lines intersect at (50, 7) so that means both plans cost $7 for 50 minutes.

B. No. Two plans DO NOT ALWAYS have the same cost for the same minutes. I made two phone plans c_1 and c_2 and set them equal. I found that x (number of minutes) has to be greater than 0 to make sense, so when you subtract the monthly fee and the slopes (cost per minute) you have to have positive values. This ONLY happens when plan 1 has the lower monthly charge and the higher cost per minute.

$$c_1 = m_1 x + b_1$$

$$C_2 = m_2 x + b_2$$

$$m_1 x + b_1 = m_2 x + b_2$$

$$m_1 x - m_2 x = b_2 - b_1$$

$$x(m_1 - m_2) = b_2 - b_1$$

$$x = \frac{b_2 - b_1}{m_1 - m_2}$$

$x \geq 0$ For the 2 phone bills to have a shared cost.

This will only occur when both $b_2 - b_1 \geq 0$ and $m_1 - m_2 > 0$.

Therefore, the plan w/ lower monthly cost must have higher per minute charge.

Figure 4.19. Student solution E to the Calling Plans task

Additionally, the conclusion conflicts with the argument, and the constraints on the monthly fee variables change throughout the argument. The solution restricts the difference between the costs per minute ($m_1 - m_2$) rates to be positive and the difference between the monthly fees ($b_2 - b_1$) to be positive. When the differences are written symbolically in ratio form, the constraint on the difference between the monthly fees is changed to include zero from previously only including numbers greater than zero. This may see to be a slight error, but it has real implications for the problem situation. Also throughout the argument, the assumption is that plan one has the greater cost per minute and lesser monthly fee. However, the conclusion claims it does not matter which monthly fee is greater. Finally, the difference between the monthly fees could be zero as well. It does not make sense to have a zero in the denominator of a fraction, but in the problem situation it means that the two plans have the same cost per minute. So if the two plans have the same cost per minute and the same monthly plan then they would also share common total costs. In summary, even though Student E did not answer the question and exhibited errors in their reasoning, seven of the eight participants said that it was a proof.

Nathaniel and Katie recognized that the solution ignored the question, but Nathaniel was the only one to label the solution a non-proof argument. He analyzed the solution and then summarized all the issues with the solution and summed it correctly by saying, "They haven't given me at least one specific example when it wasn't true [one counterexample], because their only specific example when it's not true is saying the opposite of what they said is true, which isn't true." He realized that student E did not answer the question, and the work they did present has multiple errors.

Not only did most of the participants label the argument as a proof, they thought it was the most convincing of all five Calling Plans solutions. Several participants focused on the correct use of symbolic manipulation without discussing what the symbols represented in the problem context. Tina explained that she found student's E solution to be the most convincing. When the interviewer asked why, Tina says:

> *Cause it's algebraic. As far as being the best kind of, I mean, this is kind of without a doubt talking about just how it has to work, how the equations have to work to intersect, what has to be true for them to intersect, and they're using variables and it is very general terms. It makes more sense to me than the other one did.*

Tina mentioned that the argument is justifying when the equations intersect, and she seems to miss the point of the question is to find a situation when they do not intersect. She seems more enamored by the use of variables within equations and the manipulation of the equations than studying whether the question is being answered. Uma also thought this solution was the most convincing response. She said this type of response is what you would be trying to get your own high school students to do since it is proving for all cases. Overall, six participants praised Student E and were impressed by the use of algebraic symbols without thoroughly examining the relationship between the words, symbols and problem context. Katie and Nathaniel were the only two participants who found flaws in the argument, but Katie accepted the errors. In relating this solution back to the question it is striking that some participants found this argument convincing since it never provided a counterexample.

Solution E highlights the symbolic issue that has been discussed as a concern in research related to secondary students where they accepts arguments as proof because it includes algebraic symbols even if the argument does not make sense (e.g. Healy & Hoyles, 2001). Most participants may not have been critical of this solution since it looked sophisticated. Even though the symbolic manipulation is correct, the solution did not answer the question. This raises the question of whether or not the use of algebraic symbols was the reason for the limited scrutiny, or if the participants did not connect the symbols to the problem context as they evaluated the solution. Therefore, this student solution E raises three concerns: 1) a limited understanding of variables, 2) failure to be critical of the solution since it included algebraic symbols, and 3) not knowing that a single solution is all that is needed to prove a false statement.

4.3.4 Summary of participant growth with critiquing reasoning-and-proving solutions

Overall, the participants did reasonably well as a group in distinguishing between proof and non-proof arguments, but a few challenges emerged. The distracter solutions intentionally pressed on the participants' understanding of the argument types: generic argument, empirical argument, counterexample and rationale. Counterexamples were not directly studied and proved the most challenging arguments to analyze. Some types of empirical argument solutions were easier than others to identify. Since there were only 12 distracters and they were distributed among four argument types, it is difficult to show growth.

Four of the 12 distracters were counterexamples and they all were analyzed at the same time period, so it is not possible to discuss growth in the participants understanding of counterexamples. However, a challenges arose that would be useful to further explore. It seemed as though some participants did not understand what a counterexample means or where trying to apply the criteria of proof.

The participants also evaluated four different types of empirical arguments that were meant to cause conflict. The participants were successful with identifying example only solutions and generalizations without explanations as non-proofs. However, it was more difficult for a third of the participants to identify empirical arguments when a generalization was reached using an algebraic method. Two participants accepted the algebraic procedure as prior knowledge without questioning the appropriateness of the method. Finally, seven out of nine participants confused an empirical argument with a generic example. A generic argument explains a general situation using a specific example. However, using a specific example to explain a general formula is an empirical argument. Since there were only four empirical arguments and they represented different types, it was difficult to identify growth overtime.

There are also two rationale arguments in the set of distracters. Most participants were successful with identifying unwarranted claims, but prior knowledge seemed to complicate the issue. Some participants said that it might be possible that they already learned a particular mathematical truth so it would be acceptable to state the claim. Since we are not the teacher, how do know what prior knowledge anyone student might know? The confusion with identifying prior knowledge was rectified with the class agreeing that if a claim is made in an argument that is not justified then it must state it was previously proven. While this was agreed upon while analyzing the "O + O = E" student work and there were no other rationales to evaluate, it is not possible to know if the participants improved in this regard.

While it may be acceptable to claim that participants were successful at identifying generic arguments, it is important to note that there were only four of this type (three of the four generic arguments were labeled distracters). The participants, who disagreed with how the CORP materials development team categorized the generic arguments, were concerned about how terms were defined or other legitimate disagreements. During the analysis of the first two generic argument solutions, only one participant labeled it as a non-proof based on being generalized from examples. However, after she learned the definition of a generic argument she labeled the last two generic argument solutions as proof. Therefore, eight of the nine participants were comfortable identifying generic arguments as proof from the start, and the one who was not came to accept them.

Finally, the participants were accepting of a variety of representations including the use of diagrams and narrative language, but seem to be too accepting of solutions that include algebraic symbols. The use of diagrams or narrative language was not distracting, and only one student during the first interview said that she thought proofs needed to be more formal. Most accepted narrative language and diagrams prior to the start of the course, and they all came to accept multiple representations by the end of the course. Therefore, the only concern with respect to representation is that seven of the nine participants seemed to be less critical of arguments that included symbolic manipulation.

4.4 PRE-SERVICE TEACHERS ABILITY TO SELECT AND OR MODIFY REASONING-AND-PROVING TASKS

The results in this section are in response to the fourth research question:

> 4. *To what extent do pre-service teachers improve their ability to select and or modify reasoning-and-proving tasks for students over the duration of a course focused on reasoning-and-proving and during their first year in the classroom?*

The participants were required to select two reasoning-and-proving tasks, one of which needed to be from a secondary textbook. A timeline is presented in figure 4.20 to better understand the two course assignments along with the process in which the tasks were selected and modified throughout the course. The development of task assignment one is listed below the timeline, and highlights two instances during the course where the participants were provided an opportunity to gain feedback with modifying their initially selected task. The second task assignment was an individual activity in which they selected a task from their curriculum and modified it to include reasoning-and-proving. After the course, participants were asked to select and bring a reasoning-and-proving task to the third interview. Finally, two participants with teaching positions collected reasoning-and-proving tasks they implemented as first year teachers.

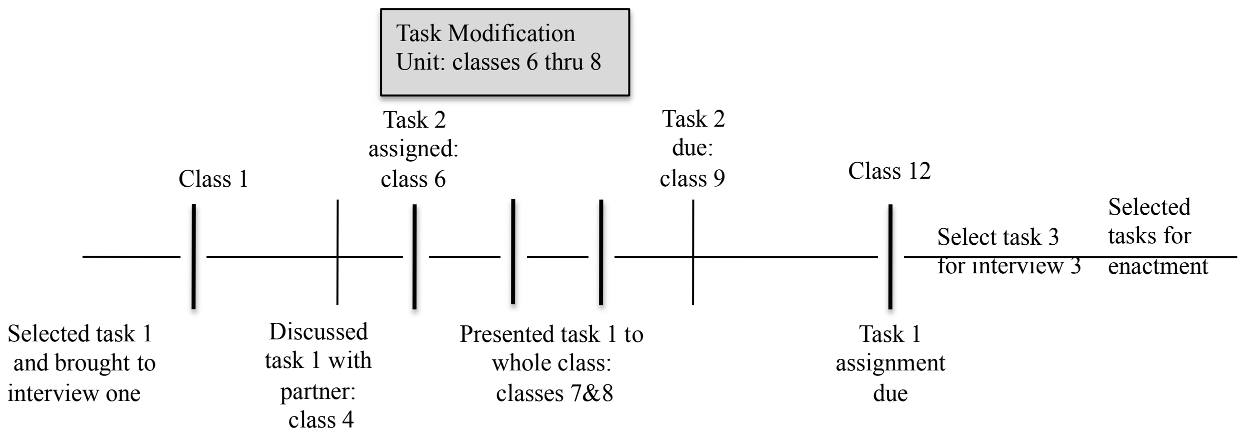


Figure 4.20. Timeline of task selection and modification

The participants were asked to bring a reasoning-and-proving task to the first interview without knowing what reasoning-and-proving meant, and they also did not know that they would be using the task as a course assignment. Eight of the nine participants brought a task to the first interview. After the participants read the Reasoning-and-Proving Framework article in class three (Stylianides, 2010) and learned that they their selected task would need to be modified to include student opportunities to reason-and-prove (class 4), they were given the option of choosing a new task. For the second assignment, the participants were required to select a task from their curricula (class 6) that had the 'potential' to be modified to include reasoning-and-proving. For interview three, the participants were asked to select and or modify a task so that students would have an opportunity to reason-and-prove. So the tasks selected for assignment one and two are similar in that the participants knew that the tasks would be modified to include reasoning-and-proving and different from the task selected for interview three since it was expected to look more like the modified versions of the tasks used in the course assignments.

This research combines the tasks selected for assignments one and two as one group of initially selected tasks, which were then modified to include reasoning-and-proving. The modified versions are coded to compare the initially selected versus the modified versions. The interview three tasks are analyzed as a group to learn if the participants were able to sustain their skill with selecting and or modifying reasoning-and-proving tasks. Finally, the 10 tasks the two participants chose to implement with their students in the 2011-2012 school year are discussed to determine what the teachers appeared to have learned about selecting reasoning-and-proving tasks.

All of the participants completed the two required course assignments. Five participants (Tanya, Uma, Karen, Katie, and Katherine) brought a task to the third interview. In addition, Karen and Katie submitted five tasks each that they implemented as first year teachers (as shown in Table 4.9)

The participants were provided the option of modifying the initial task they selected or choose an entirely new task. The numbers in the parentheses represents the number of times the participant modified their selected task. Six participants selected a task and made modifications on two separate occasions (Nathaniel, Tina, Lucy Uma, Brittany, and Katherine). Karen modified her original task once. Tanya chose to select a second task and then modified the new task once. Katie selected a second task, modified it then made a decision to choose a third reasoning-and-proving task.

Table 4.9. Total number of reasoning-and-proving tasks each participant selected

Participant	Assignment 1 (modified)	Assignment 2	Interview 3	Selected & implemented as 1st year teachers
Nathaniel	1 (2)	1	0	0
Tanya	2 (1)	1	1	0
Karen	1 (1)	1	1	5
Tina	1 (2)	1	0	0
Lucy	1 (2)	1	0	0
Uma	1 (2)	1	1	0
Brittany	1 (2)	1	0	0
Katie	3	1	1	5
Katherine	1 (2)	1	1	0

4.4.1 Ability to modify tasks to include reasoning-and-proving opportunities

The twenty[10] initially selected course tasks were coded using a two dimensional matrix that combines the Reasoning-and-Proving Framework and Task Analysis Guide (TAG) (as shown in table 4.10). The tasks were determined to be either high level (procedures with connections or doing mathematics) or low level (procedures without connections or memorization) along the TAG dimension and coded as either a call for an argument (proof or non-proof) or a requirement to make a generalization (identify patterns or make a conjecture). It is also possible that a task was not a reasoning-and-proving task.

[10] Two for each participant and one extra for Katie and Tanya.

Table 4.10. Results of initially chosen reasoning-and-proving tasks

	Make a generalization	Provide an argument	Not a reasoning-and-proving task
Low-level task	0	1	0
High-level task	11	5	3

The results show that one of the chosen tasks was low level and required a proof. Just over half of the 20 selected tasks were high-level and required a generalization. One fourth of the tasks were high level and directly asked students to justify specific cases or explain why a conjecture is always true. So eighty percent of the selected tasks were high-level reasoning-and-proving tasks. However, three[11] of the "high-level provide an argument" tasks were not used for assignment one. Therefore, while the goal was for participants to modify tasks to be "high-level provide an argument" type tasks, additional factors contributed to how and why tasks were modified and or discarded.

It is important to note that these selected tasks were not expected to be of a particular type. It is possible that students purposefully did not chose a proof task since they may have thought it would not be modifiable. Also, the CORP materials includes the case of Nancy Edwards, which highlights the modification a "low-level make a generalization" task to be a "high-level provide an argument" task. Therefore, the participants were left to determine what a reasoning-and-proving task with potential meant.

[11] Katie abandoned two and Tanya abandoned one of the tasks.

Growth was detected by comparing the eighteen[12] originally identified tasks to the final form of the two course assignments (as shown in table 4.11) and to the task selected for the third interview, and ultimately the tasks selected and used in the classroom. The 18 initially selected problems are listed in the first column, and the second row lists all six possible outcomes of task types in order from least to most sophisticated. For instance, the 11 high-level "make a generalization" tasks were modified in three different ways where nine became high-level "provide an argument" (column 7). One task became a low level provide an argument (column four). At least one task from each of the four original categories did not change task type after modification and this is shown in the table where the numbers form a diagonal line. No reasoning-and-proving tasks were altered to become non-reasoning-and-proving problems. Overall, 11 tasks were modified to improve their sophistication level (above the diagonal line), six stayed the same (those along the diagonal line) and one task was modified to where the cognitive demand was lowered.

Table 4.11. The 18 selected and modified reasoning-and-proving course tasks

Initially Selected	Modified Course Assignments Tasks					
	L-L Non-R&P	L-L Make a G	L-L Provide an Argument	H-L Non-R&P	H-L Make a G	H-L Provide an Argument
(1) L-L Provide an Argument			1			
(3) H-L Non-R&P				1	1	1

[12] Katie's third task is labeled a modified version of her second task. The two tasks Katie and Tanya selected and abandoned are not included.

(11) H-L Make a G			1		1	9
(3) H-L Provide an Argument						3

In general, the participants learned to apply the criteria discussed during the course to modify tasks. During the task modification unit in the course, several activities led to the development of a set of principles as shown in figure 4.21. The idea was that asking the types of questions listed in the principles would provide opportunities for students to reason-and-prove. The participants learned to modify tasks to include these principles.

Task modification principles

- Scaffolding
    - Remove scaffolding (to increase number of solution paths)
    - Add scaffolding
        - Organize thinking
        - Multiple Entry points: access
- Ask Why? Or Why not?
- Connect Representations
- Explore patterns – make observations
- Students produce conjectures

Figure 4.21. Course developed principles for reasoning-and-proving task modification

Most (13/18) of the modified written tasks provide students with opportunities to provide high-level arguments. Some revisions removed scaffolding to focus on argumentation, others added questions so that students could review multiple examples before making a generalization. An interesting finding is that the number of examples requested before asking students to provide an argument varied between one and four.

Many modified tasks included asking why? or why not? As a follow-up question to calling for a generalization, many revised tasks included the question: “How do you know?” So instead of asking why is your generalization true or write a proof, the participants overwhelming opted to ask how do you know that your conjecture will be true for any possible situation. However, simply adding this question to a task does not make it high-level or a proof task. There needs to be a conjecture in the problem for students to justify and the task must have multiple solution paths or opportunities to make connections. Therefore, not only were most (8/9) of the participants able to select or modify reasoning-and-proving tasks during the class, they applied multiple principles creating opportunities for students to reason-and-prove.

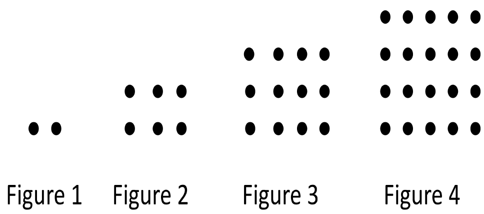
The next section explains how the tasks were changed to align with the different categories including the modification principles. Tasks in each of the four initial categories are discussed in juxtaposition with how they were altered.

4.4.1.1 High-level make a generalization task

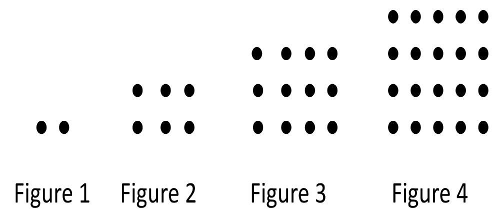
The 11 high-level “make a generalization” tasks were selected for both the first task (n=4) and the second task (n =7) selection assignment. The 11 tasks were grouped into three categorizes: 1) pattern tasks (n =4), 2) examination of cases tasks (n =5), and 3) extrapolate general features from a single example tasks (n =2). Four of the eleven are pattern tasks and the other five examination of cases problems (second category) prompted students to analyze a set of examples, polygons, or numbers. The final two tasks in the third group focused on a single case, and asked general questions from the provided example. The distinction between “make a generalization” and “provide an argument” is that provide an argument tasks must explicitly require an explanation for why a specific or general case always works. In other words, explain a rule (generalization) was coded different from explaining why a rule is true for all cases (provide an argument). One of each of the three types of high-level "make a generalization" tasks along with how they were modified are shared.

While the four prototypical-pattern tasks varied, they each followed a common modification structure: 1) extend pattern past what is given, 2) explain or describe a figure without drawing it, 3) make a generalization and explain why it is always true (see to the rectangular dot pattern Karen selected and modified in figure 4.22 as an example). Karen made several changes to modify her task from "make a generalization" to "provide an argument." In other words, she did not simply write prove your generalization as a sixth question. The first set of questions focus on the number of dots, how the numbers are changing, and finally to write an equation based on the numbers extracted. Her modification draws attention to how the organization of the dots are changing and making connections with the figure number. For instance, her modified question (d) does not simply ask for the number of dots for the 10th figure, since she is more interested with how the student is thinking about the shape. These changes support students with answering part two of question (e) about justifying how students know their equation will be true for any counting number.

Initially Selected Version	Modified Version



a. What are the first four rectangular numbers?
b. Find the next two rectangular numbers.
c. Describe the pattern of change from one rectangular number to the next.
d. Predict the 7th and 8th rectangular numbers.
e. Write an equation for the *n*th rectangular number *r*.



a. Write down everything you observe about the pattern.
b. What are the first four rectangular numbers?
c. How do the numeric values relate to the picture?
d. Describe the picture of the 10th rectangular number.
e. Use words, diagrams, or symbols to generalize the pattern. How do you know your generalization is true?

Figure 4.22. Karen's initial task and modified version completed her second course assignment

Of the five "examine cases tasks" four were modified to require an argument. The final task did not initially require students to generalize the situation. Two of these tasks pertained to the exponent rules, two others related to the interior angle measure of polygons, and the problem related to exponential decay.

Tina's modification of her exponent rule task, which stayed high-level and required an argument, is shared in figure 4.23. The original exponents task that Tina selected is labeled high-level "make a generalization" since it prompts for an explanation of the conjecture. A judgment is made here that *explain* does not mean that same as "provide an argument." In this case, explain means to tell why you choose positive or negative, and citing the four examples is considered sufficient evidence.

The modified version includes the same examples, but removes the "either or" conjecture. Students are open to make a variety of observations before focusing on negative numbers raised to an odd and even whole number exponent. Tina's part (b) for both questions three and four presses students to think beyond the provided examples to make an argument for both cases. The addition of these two questions account for why the modified task is labeled "provide an argument."

Initially Selected Version

1. Simplify each expression

   $(-2)^2$ $(-2)^3$ $(-2)^4$ $(-2)^5$

   $(-3)^2$ $(-3)^3$ $(-3)^4$ $(-3)^5$
2. Make a conjecture: Do you think a negative number raised to an even power will be positive or negative? Explain
3. Do you think a negative number raised to an odd power will be positive or negative? Explain

Modified Version

1) Solve the following examples.

   $(-2)^2$ = _____ $(-2)^3$ = _____ $(-2)^4$ = _____ $(-2)^5$ = _____

   $(-3)^2$ = _____ $(-3)^3$ = _____ $(-3)^4$ = _____ $(-3)^5$ = _____
2) Make some observations about any patterns that you notice.
3) a. Using what you notice about the examples above, make a conjecture about negative numbers to an even power.

   b. How do you know that this will be true for all negative numbers?
4) a. Using what you notice about the examples above, make a conjecture about negative numbers to an odd power.

   b. How do you know that this will be true for all negative numbers?

Figure 4.23. Tina's exponent task that is high-level and make a generalization

The final two tasks ask students to make generalizations from a single case, and Lucy identified and modified both. In one task, after exploring the single example, students are expected to generalize convergence and divergence for the area for a general situation ($y = x^n$). Parallel lines problem is shared in figure 4.24 to further explain the case in which a task promotes generalizing and proving a situation from a single example. The initially selected version is labeled "make a generalization" since it asks students to generalize the relationship between the slope and y-intercept of two equations that are parallel from a single case. Students are not asked to consider additional cases or justify why their conjecture is true for any pair of parallel lines.

Lucy modified the task to have students start to consider the possibilities of any pair of linear equations before focusing on the particular system. The task requires that students solve the pair of equations in multiple ways including a graphical representation. After examining this specific case, questions prompt students to think about any pair of parallel lines. The wording of the original version caters to students making a conjecture after exploring a single example. Lucy's modified version is also centered on the same case, but she words the questions so that students are expected to provide support or a counterexample to the general case.

Initially Selected Version

Parallel lines problem: Show that the graphs of 3x – 2y = 6 and 6x – 4y = 18 must be parallel lines by solving each equation for y. What is the slope and y-intercept for each line? What does this mean? If a linear system is inconsistent, what must be true about the slopes and y-intercepts for the system's graphs?

Modified Version

1. What do the solutions to a system of equations represent graphically?
2. Solve the following system:

   3x – 2y = 6
   6x – 4y = 18

   Definition: A linear system with no solution is called <u>inconsistent</u>.
3. Interpret your solution in terms of another mathematical representation.
4. Show that the lines given by the following two equations are parallel.
5. Explain why your solution to #4 proves that the lines are parallel and why this makes sense.
6. Can the solution to a system of equations of lines, which are not parallel, ever be inconsistent? If yes, give an example to verify this statement. If no, explain why not.
7. Can lines that are parallel ever have one or more solutions? If yes, give an example to verify this statement. If no, explain why not.

Figure 4.24. Lucy's selected and modified parallel lines problem

The three examples discussed in this section are different with respect to the amount of scaffolding each task provides even though each task was modified away from "make a generalization" to "provide and argument." For example, Karen's modification (figure 4.21) included question (d) to support students with moving from concrete provided examples to explaining the general structure of the pattern. Lannin (2005) suggests that this type of scaffolding supports students with constructing a generic argument. Tina (figure 4.22) asked students to examine a few cases, but never had them consider examples beyond what she presented. Instead, students are expected to move from the provided set of examples to justifying why what they observe will always be true. Finally, Lucy's (figure 4.23) modification provides less scaffolding since students are only given one example to explore.

4.4.1.2 High-level provide an argument tasks

Five of the initially selected 20 tasks for one of the two course assignments included opportunities for students to provide an argument. Three of the five proof tasks, however, were discarded since the participants were uncomfortable with how they might engage students in solving the tasks. Of the two remaining proof tasks only one explicitly called for a proof. The other promoted a non-proof argument.

Katie selected two of the five high-level "provide an argument" tasks and she did not modify either task for the two course assignments. After trying to modify one task (shown in figure 4.25) she came to the realization that it might be too difficult for students to access. Her second task provides multiple solution paths making it more accessible to secondary students, which was her rationale for discarding the first problem. Therefore, even though both are identified as high-level "providing an argument," Katie considered how students could solve the task including possible representations as additional principles for selecting her final task assignment.

Initially Selected Version (discarded)
Multiply 4 consecutive positive integers and add 1 to the product. What kind of number do you get? Will this always happen? If you think so, prove it.

Newly Selected Version
Pick any positive integer. Add 2 to it. Take the product of this number and your original number and add 1. Make a conjecture about the resulting number. Try more examples if you need help conjecturing. Will your conjecture always be true? Find a counting number that does not work OR show why your conjecture always works.

Figure 4.25. Katie's assignment one proof tasks

Nathaniel also modified a high-level "provide an argument" task, which prompted students to prove that the formula for the area of any triangle is A = ½ bh. Even though the task started as a high-level providing an argument task, Nathaniel still considered ways to modify the problem (as shown in figure 4.26). Nathaniel considered the principle that students may not know how to start the original problem. The modified version includes much more scaffolding with the requirement to examine four specific examples (inductive reasoning) to provide students access and for students to make an explicit connection between the number of unit squares and the formula. The third question moves away from examining specific cases to thinking about the relationship for any triangle. Therefore, Nathaniel included scaffolding questions to modify the initial task from strictly a proof task to one that includes a broad range of reasoning-and-proving activities.

Initially Selected Version *Area of a Triangle Task:* You have always been told that the area of a triangle can be obtained from the formula (1/2 bh). • But how do we know that is always true? Will this formula really work for any triangle? Prove that this formula will provide the area for *any* triangle.
Modified Version *Area of a Triangle Task:* 1. The area formula for a triangle is given by (1/2/ b h), but where does this formula come from? Investigate the origin of this formula by filling in the table below (*see italic*). • First, identify the base and height for each triangle. • In the third column calculate area by using the formula. • In the last column, find an approximation for area by determining the number of squares inside each triangle. *(Provides a table for students to complete: base, height, area, # of interior squares)* *(Provides four triangles on a grid for student to use to complete table)* 2. Draw a new triangle in the grid below and explain how its area is connected to the formula: ½ bh. 3. Explain why the formula, ½ bh, is always the same as the number of squares inside of a triangle? 4. Are you convinced that the formula will work for all triangles? Why or why not?

Figure 4.26. Nathaniel's modified Area of Triangle Task

4.4.1.3 Low-level reasoning-and-proving tasks

Two tasks were modified to be a low-level 'provide an argument," and one of the two started in the same category. One of Katherine's tasks that included rules for exponents started as a high-level "make a generalization," and was modified to have students provide a justification for two laws of exponents rules. Brittany did modify her task (as shown in figure 4.27), but the changes were not enough to increase the cognitive demand.

The solution method is considered a derivation of the formula and is low-level since it only promotes use of an algebraic procedure without making connections to why or how the constants relate to the y coordinate of the vertex. Brittany's revisions to her task attempts to provide access through inductive reasoning after students are asked to consider what they know about parabolas. However, the questions do not scaffold students toward providing an argument.

Students will need to create their own quadratic equation examples and graphs, but it seems unrealistic to think that students could make connections between the coefficients that they choose in their examples and the y coordinate of the vertex in the graphs especially given the complexity of the relationship ($-(b^2 / 4a) + c$). Therefore, since the scaffolding does not support students in writing an argument or with making connections across representations, the task did not change from its original categorization.

Initially Selected Version
For the graph of $y = ax^2 + bx + c$ show that the y coordinate of the vertex is $-(b^2 / 4a) + c$.
Modified Version For the graph of $y = ax^2 + bx + c$: a) Compile a list of everything you know about parabolas. b) Graph a few parabolas and make a list of observations about how the vertex relates to the rest of the graph. c) *Prove* or *Disprove* that the y coordinate of the vertex is $-(b^2 / 4a) + c$ for all parabolas.

Figure 4.27. Brittany's parabola problem, which stayed a low-level provide an argument task

4.4.1.4 High-level non-reasoning-and-proving

Three of the originally selected tasks were not reasoning-and-proving tasks. One was modified to become a "high-level make a generalization," one remained as non-reasoning-and-proving, and one was modified to become a high-level 'provide an argument." So none are representative of the group, but the two that changed are different from the kinds of tasks previously shared.

Tanya's task was modified from a non-reasoning-and-proving to a "make a generalization," and the cognitive demand of both tasks stayed at a high-level (as shown in figure 4.28). Tanya's original task asks students to apply their understanding of parabolas to find a specific example, but it does not provide an opportunity for students to reason-and-prove. The modified version is not a typical "make a conjecture problem." Students are not asked to find a formula or explicitly state a conjecture. However, the second question asks students to consider multiple cases. Students could generalize the situation from exploring multiple cases, but examining cases and explaining what is noticed constitutes identifying a pattern, which is part of "make a generalization."

Initially Selected Version
Is it possible to make a parabola that lies only in quadrants II, III, and IV? If so, write an equation for such a parabola. If not, say why not.
Modified Version
1. Is it possible to make a parabola that lies only in quadrants II, III, and IV? If so, write an equation for such a parabola and explain how you know it only lies in these quadrants. If not, say why not. 2. How many such parabolas exist? Explain your answer.

Figure 4.28. Tanya's parabola problem

Katherine's task was modified from non-reasoning-proving to "make a generalization" (as shown in figure 4.29). The original task is label high-level based on the second question in which students are asked to make connections between the balloon arch, the graph, and the equation. However, they are not asked to identify a pattern, make a conjecture or provide an argument. The modification could be considered low-level. Students that are unable to solve the equation, make a take of values, or graph the equation to find the x-intercepts or vertex will not be able to start the problem. However, students are not asked to follow a particular method and multiple methods are possible. Katherine removed the diagram and changed the questions. The mathematical focus changed from finding possible x values for the situation to discussing a general connection between x-intercepts and the x value of the vertex for any quadratic function. Students are not prompted to explore additional examples, but as was previously discussed the task could be an attempt to remove examples so that students learn to generate their own. Finally, the conversation around this task could be broader than Katherine intended. Not all quadratic functions intersect with the x–axis leading to lots of possible solutions based on the constraints different students may place on parabolas and their understanding of functions.

Initially Selected Version An arch of balloons decorates the stage at a high school graduation. The balloons are tied to a frame. The shape of the frame can be modeled by the equation $y = -(1/4)x^2 + 3x$ where x and y are measured in feet. a. Make a table that shows the height of the balloon arch for x = 0, 2, 5, 8, and 11 feet. b. For what additional values of x does the equation make sense? *Explain*. c. At approximately what distance from the left end does the arch reach a height of 9 feet? Check your answer algebraically. *A diagram showing the arch and indicating that the maximum is 9ft is included.*
Modified Version An arch of balloons decorates the stage at a high school graduation. The balloons are tied to a frame. The shape of the frame can be modeled by the equation $y = -(1/4)x^2 + 3x$ where x and y are measured in feet. 1. What are the x-intercepts of the function? 2. What is the vertex of the function? 3. How do the x-values of the x-intercepts and the vertex relate? 4. Will this be true for all quadratic functions? Explain how you know. *No diagram is provided.*

Figure 4.29. Katherine's quadratic function problem

4.4.1.5 Summary of tasks selected and modified for the two course assignments

All nine participants selected two tasks each for which they saw potential to modify to include reasoning-and-proving (as shown in table 4.13). The participants' names are listed in the first column, and the next two columns in order represent the initially selected and modified coding for each of the two course tasks. The codes were abbreviated to cut back on the amount of text in each cell. The 1) and 2) used to represent the first and second course assignments. The shaded rows are included to help with distinguishing between the participants.

Table 4.12. Participants' skill with selecting and modifying reasoning-and-proving tasks

	Initially selected versions	**Modified versions**
Nathaniel	1) H-L argument	1) H-L argument
	2) H-L generalization	2) H-L argument
Tanya	1) H-L not reasoning-and-proving	1) H-L generalization
	2) H-L generalization	2) H-L argument
Karen	1) H-L generalization	1) H-L argument
	2) H-L generalization	2) H-L argument
Tina	1) H-L generalization	1) H-L argument
	2) H-L generalization	2) H-L argument
Lucy	1) H-L generalization	1) H-L argument
	2) H-L generalization	2) H-L argument
Uma	1) H-L argument	1) H-L argument
	2) H-L generalization	2) H-L generalization
Brittany	1) H-L not reasoning-and-proving	1) H-L not reasoning-and-proving
	2) L-L argument	2) L-L argument
Katie	1) H-L argument	1) H-L argument
	2) H-L generalization	2) H-L argument
Katherine	1) H-L generalization	1) L-L argument
	2) H-L not reasoning-and-proving	2) H-L argument

Five participants' (Nathaniel, Karen, Tina, Lucy, and Katie) tasks were all modified to be high-level proof tasks. Both of Uma's and Brittany's tasks were modified, but stayed in the same categories. Tanya's non-reasoning-and-proving task became a "make a generalization" and her second assignment “make a generalization” task was modified to “provide an argument.” Katherine’s “make a generalization” problem became a low-level “provide an argument”, and her non-reasoning-and-proving problem became a high-level “provide an argument” task.

4.4.2 Five selected tasks for interview three

All nine participants were asked to bring a reasoning-and-proving task to the third and final interview, and five participants (Tanya, Karen, Uma, Katie, Katherine) brought a task. All five selected tasks were high-level. Four (Tanya, Karen, Uma, Katherine) are high-level “provide an argument” and the fifth (Katie) is high-level non-reasoning-and-proving. The point in asking them to bring a task to the third interview was to begin to understand if they could continue to select and or modify appropriate reasoning-and-proving tasks.

None of the third interview tasks were selected from a textbook, and only one task was slightly modified. Two of the tasks were participant designed pattern tasks, and the other three tasks were selected from a resource. Tanya and Katherine both designed pattern tasks that follow a similar sequence of questions as Karen’s rectangular number pattern task (figure 4.21) shared in the previous section. Students were asked to examine three or four figures before explaining a figure that is too big to draw. Tanya modified a similar linear task for the second course assignment, but this was Katherine’s only pattern task.

After choosing a pattern task for each of her first two assignments, Karen chose a different problem for the third interview. Instead she selected a task called the Blocks Task (as shown in figure 4.30) from a previous course as a graduate student and added the question: “How do you know your answer is correct?” The Blocks task allows students to either identify a particular or general solution and the follow-up question Karen added requires students to develop an argument for their solution to the first question. This question promotes the use of a manipulative and helps students to make connections between factors and multiples. In other words the question provides students an opportunity to construct an argument through the use of tools and supports students in exploring typical mathematics content.

Blocks Task

Yolanda was telling her brother Damian about what she did in math class.
Yolanda said, "Damian, I used blocks in my math class today. When I grouped the blocks in groups of 2, I had 1 block left over. When I grouped the blocks in groups of 3, I had 1 block left over. When I groped the blocks in groups of 4, I still had 1 block left over."
Damian asked, "How many blocks did you have?"
What was Yolanda's answer to her bother's question?

Karen added: How/why do you know your answer is correct?

Figure 4.30. Karen's blocks task

The final two tasks were not selected from a conventional curriculum either. Uma choose a number theory task in which the solution is a counterexample. A conjecture is provided with a series of examples and students are expected to decide if it is true. Katie chose the only task that was non-reasoning-and-proving (as shown in figure 4.31). Katie's squares task is high-level since it can be solved in many different ways where the sides of the squares can be labeled with generic numbers or variables. An accurate solution would require students to apply the Pythagorean theorem.

Squares, Shaded Area

The figure at the right consists of squares and isosceles triangles.

What percent of the entire figure is shaded?

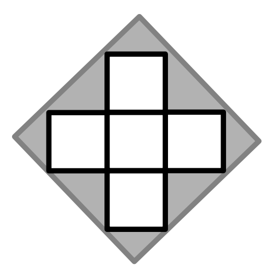


Figure 4.31. Katie’s squares and shaded area problem

4.4.3 Ten implemented classroom reasoning-and-proving tasks

Katie and Karen each selected and implemented five high-level reasoning-and-proving problems during their first year as secondary teachers, and all 10 tasks were high-level “provide an argument.” The two teachers capitalized on modifying tasks from both their curriculum and the CORP course. While most of the tasks were modified, in this analysis only the enacted tasks are discussed and not how the task may have appeared in a curricula resource.

4.4.3.1 Karen’s enacted proof lessons

Karen enacted the tasks in her high school geometry classroom. While it may seem as though there would be many opportunities to engage students in proof in the geometry curricula, Karen still choose tasks outside the content area and modified the tasks she selected from her textbook. She implemented two pattern tasks, two tasks from her curricula materials, and one task related to her content from an outside resource. While it is not certain these are the only proof tasks she enacted during the 2011-2012 school year, the five she submitted show that she did provide opportunities for students write proofs.

Karen’s first three reasoning-and-proving tasks were related to her geometry curriculum. She focused on wanting students to understand that a proof is mathematical argument that can take on many different forms in which they ultimately need to convince others of the truth. This is evident throughout each of her first three tasks.

The first task asked students to write an argument about how to divide a square into four equal parts. Students were expected to explain and convince their classmates why the parts of their square that they drew were equal. This activity provided her students an opportunity to construct an argument that included the connection of words and a diagram while applying prior knowledge about squares and area.

Karen's second modification required that students write a proof of a situation using two different forms since she expected all students to follow a two-column form for at least one method. While the solution was basic in that it only required two steps, Karen's point was for students to explain the proof using two forms.

Karen modified the 3rd task to separate the conjecture from the argument and choose to include 'proof or explanation' so that students do not feel constrained to producing a two-column structure. First students were asked to apply prior knowledge about a transversal and two parallel lines to make a conjecture about same side exterior lines before justifying their conjecture. Therefore, these three tasks Karen enacted were connected to her curricula, and she made similar modifications across the problems to support students in learning that a proof can assume many different forms.

After the third task, Karen implemented two pattern tasks. As a participant in the reasoning-and-proving course Karen modified two pattern tasks and choose to implement one of them with her geometry students: Pool Border problem. The other pattern task was the "S" pattern problem she solved in a previous graduate course. She implemented the Pool Border problem as she previously modified it. However, she modified the "S" pattern task to include reasoning-and-proving which emulated the structure she and others developed in the course: 1) draw next two figures, 2) explain what the 50th figure would look like, and 3) generalize and prove for any size figure. So it is evident that not only was Karen comfortable and successful with modifying pattern tasks during the course, she was able to modify similar tasks as a classroom teacher.

4.4.3.2 Katie's enacted proof lessons

Katie taught a mixed 11^{th} / 12^{th} grade pre-calculus course, which she= labeled as remedial. She modified four of her five tasks to extend toward proof, make connections between representations, and or to "plant a seed of doubt." Three tasks were selected from the CORP materials, one modified from her textbook, and the fifth was from her methods course. Only two of Katie's five tasks related to the content she was teaching. Even though proof tasks are typically not implemented outside high school geometry, Katie was motivated to provide at least five opportunities for her students to engage in these practices.

Katie explained that she implemented a modified version of the "O + O = E" task from the course after she finished a chapter in her curricula on unit circle. She modified the task so that students would complete a conjecture and justify the statement or find a counterexample, which she also labeled as "plant a seed of doubt." She implemented this task just before the December holiday break. After a unit on probability, Katie enacted a pattern task that involved finding the perimeter of a hexagon pattern. She modified it to ask students if they are sure their formula will always work. Prior to a unit on sequences and series, Katie implemented the Sticky Gum problem without any modifications. Therefore, three of Katie's proof tasks were selected or modified to provide her pre-calculus students an opportunity to provide arguments, but the problems were disconnected from her course curriculum.

During the unit on sequences and series, Katie included two of the exploring and explaining visual proof tasks as a single activity that she solved as a student in the CORP course. She modified the geometric series that were originally accompanied with a diagram to connect with the terminology used in her textbook. Also, while engaged in a unit on the binomial theorem, she identified a triangular number pattern that she modified and implemented. Her modification promoted a connection between the number of dots in each triangle and the figure number. Therefore, Katie was able to include proof tasks that connected to her curriculum when the content related to tasks she already solved or recognized a pattern task in her textbook, which she knew how to modify.

4.4.4 Summary of selecting and modifying reasoning-and-proving tasks

Two of the required assignments for CORP course, was for the participants to select a task from their curricula and one from any other resource. After engaging in several activities the participants identified general modification principles in which they used. When asked to look through curriculum resources to select a task with potential, several participants focused on pattern tasks. The inductive nature of the problems align well with the reasoning-and-proving framework since students can look for patterns, make a generalization, and explain why the formula works for all cases. The participants began to recognize the inductive structure to modify tasks to include a set of examples, or used a single example in which students were expected to extrapolate generality. For instance, Nathaniel added the requirement for participants to first find the area of four triangles before explaining why the formula (A=1/2(bh) is always true. Additionally, Katherine modified her parabola problem for students to explore one quadratic function before discussing the relationship between x-intercepts and the x value of the vertex. It is unknown if those that provided modified tasks that only included one example expected students to generate more to justify an observed relationship or if they thought students could recognize generality from a single case, typical of deductive proof tasks in Euclidean geometry courses.

An interesting finding was that none of the required 18 course tasks the participants selected were from a geometry textbook. However, Karen taught geometry her first year and did not share any proof tasks that promoted students understanding of geometric concepts through inductive reasoning. For example, one task presented a pair of parallel lines with a transversal. Students were not asked to explore several cases in which the transversal cut the lines at various angle measurement, nor was the task altered to ask students to investigate what happens if the parallel lines were close together or further apart. Instead they were given a single example and were expected to prove the situation using deductive reasoning. In other words, Karen modified geometry tasks so that students would produce multiple forms instead of modifying them to resemble pattern tasks so that students could make sense of the definitions through inductive reasoning.

Overall the participants proved capable with identifying tasks to modify that include an inductive pattern. Also many were able to select proof tasks from alternative resources that were accessible and included students to make a conjecture and write a proof, which were mostly in the number theory content. Two teachers selected five high-level "provide an argument" tasks each and enacted them with their students, and the content of the tasks align with the course curricula five out of ten times.

5.0 CHAPTER 5: DISCUSSION

In this chapter, the findings reported in chapter 4 are discussed more broadly to explain the implications for teacher learning, mathematics education research, and the design of professional development curricular materials. The first section integrates relevant research while summarizing the results of this study with respect to the prospective teachers' learning about proof. Next, comparisons are made between what the participants said and how they actually completed the course and interview tasks. Additionally, the participants' ability to construct proofs will be contrasted against their skill with validating arguments. Finally, the chapter provides a conclusion and directions for future research.

5.1 IMPORTANCE OF STUDY: EXPLAINING THE RESULTS

This design research study provided the participants opportunities to learn about proof. The course expanded participants' conceptions of proof and identified the challenges prospective teachers face when they engaged in proof activities, including selecting and/or modifying reasoning-and-proving tasks. Current research provides information on prospective and in-service teachers limited views of proof (e.g. Knuth, 2002a, 2002b, Solomon, 2006, Smith, 2006), their inability to distinguish between proof and non-proof arguments (Knuth, 2002a, Morris, 2002; Selden & Selden, 2003), their lack of skill in constructing a proof (e.g. Moore, 1994; Morris, 2002; Recio & Godino, 2001; Weber, 2001), and their pedagogical challenges with supporting students in producing a proof (Bieda, 2010; Edwards & Ward, 2004; Martin et al. 2005; Smith, 2006). The work to date has been useful with identifying limitations in teachers' knowledge and practice, but little is known regarding how to address these limitations. This research study aimed to gain insights into what prospective teachers understand, believe, and struggle to learn as they engage in a course designed to improve their knowledge and ability to enact reasoning-and-proving tasks with students. The results suggest participants did expand their conception of proof and important insights were gained as they were asked to construct arguments, analyze student solutions, and select and/or modify reasoning-and-proving tasks. The next four sections will discuss the results of each research question in connection with the existing research on proof in secondary mathematics education.

5.1.1 Expanded conception of proof

As a group, the participants changed their conceptions of proof through the engagement in various course activities. While the four categories pertaining to the conception of proof (criteria, purpose, equity, and opportunity) were analyzed individually, there seems to be obvious connections across them. In other words, as a participant expanded their understanding and beliefs in one area, say purpose of proof, then it seemed to affect another conception area, such as the ability to provide more opportunities for proof. These connections will be discussed further in this section.

At the beginning of the course, most participants mentioned that a proof needed to include logical steps to show why a statement is always true. These results are similar to what

Knuth (2002b) found when he asked teachers what constitutes a proof and Smith's (2006) findings about undergraduate students understanding of what makes a proof valid. In other words, the form of the argument includes definitions and statements following from the conjecture to the conclusion. During follow-up interviews, the participants identified course activities that attributed to their expanded view of proof. Many participants specifically mentioned how the analysis of students work ("O + O = E": task 3, Appendix 3.2) broadened their view of the representation of a proof. So while a proof could include a series of statements and definitions that lead to the conclusion, they also began to accept that proofs could include diagrams and everyday language as well. In addition, they learned the importance of developing a list of commonly accepted definitions and mathematical statements; allowing the class community to keep track of what claims require further justification and which ones do not need explanation, since they were already proven and are accepted truths (Hanna, 1990). Since all of the participants expanded upon their original perception of what counts as proof, they were able to communicate new reasons for including proof in secondary mathematics.

The course expanded participants' conception of the purpose of teaching proof in secondary mathematics from three initial reasons to a total of seven at the end of the course. The participants identified most of the purposes for proof in secondary education suggested in research (Bell, 1976; de Villers, 1990; Hanna, 2000), which were also identified by the teachers in Knuth's (2002b) study. Additionally, the participants explained that they believe students should engage in proofs to develop their own mathematical authority (Harel & Sowder, 1998; Smith, 2006). Finally, the participants identified specific course activities (i.e. Case of Nancy Edwards, reading articles, etc.) as contributing to their new realization that students need to engage in proof tasks so they can learn to do them and understand what constitutes a proof. It is difficult to assign causality, but it is reasonable to believe that since the participants gained a more defined criteria of proof, they began to think about additional purposes to include it. For instance, since the participants accepted the fact that a proof is not an objective product, they also recognized that engaging students in construction could develop communication skills and build students mathematical authority. In other words, a change in the criteria of what a proof could look like may have supported the participants' thinking about the kinds of tasks students could engage in solving and how such an activity has the potential for multiple purposes during instruction.

Knuth (2002b) found that the teachers in his study possessed varying views of proof (formal versus informal), and those with a formal view did not believe proof should be included in high school mathematics except for maybe honors students. However, the teachers with an informal view of proof recognized its usefulness and applicability across all courses for all students. This was also the case in this current study. A broader understanding of proof supported the participants in recognizing how reasoning-and-proving tasks can be implemented more often in all secondary courses and with all students. The difference between Knuth's study and the current study is that the participants in the study reported herein learned that informal empirical arguments are not sufficient and they believed it is possible to hold students accountable for developing valid arguments. Therefore, the participants not only expanded their understanding of what counts as proof, they came to believe that all students can construct proofs for a variety of mathematical purposes with appropriate support.

A view that the course was not able to change for the same teacher is the belief that the purpose of secondary mathematics is to cover the school or district adopted textbook chapters and with limited number reasoning-and-proving tasks in the curricula (Thompson, Senk, Johnson, 2012), there is not enough time to fit it in. Given the girth of topics most secondary books include, the foreseen time commitment to supporting students with learning to construct arguments, and the fact that state tests did not assess student knowledge of proof, a few participants expressed reluctance to incorporate proof tasks into their future courses.

5.1.2 Constructing proofs

There existed two major reasons why participants were unsuccessful with producing a greater number of proofs: 1) a limited understanding about the meaning of variables; and 2) an inability to develop a clear general argument or one without assumptions. Some participants thought that a variable could hold multiple meanings in the same problem. This issue was most evident in the tasks that were not situated in an everyday context ("O + O = E," $N^2 + N$ is always even, Explaining Number Patterns). For example, one participant wrote n is even and followed it with writing 2n. As a reader, this could be interpreted to mean 2n represents multiples of four, even though the problem states that n is any counting number. This type of response suggests the participant did not realize that she defined the same variable in two different ways in the same problem, which is mathematically incorrect. In the same task ($N^2 + N$), two other participants defined an odd and even case for n, which is an acceptable way to use the same variable in the same problem. However, their definition of odd ($n = 2k + 1$) where k is a natural number does not allow for n to be any counting number. The second example could be an oversight and might be corrected by simply highlighting the error, but it could be a larger issue about understanding variables. The third issue is that participants introduced variables without defining them. In all, seven of the nine participants demonstrated at least two of the three defining variable issues.

Every participant initially solved each of the three non-contextual tasks ("O + O = E," $N^2 + N$ is always even, Explain Number Patterns) using algebraic symbols. However, current research has not addressed how prospective or practicing teachers use variables when constructing proofs. Findings involving secondary students, suggest that most of them do not use symbols when asked to write proofs (e.g. Bell, 1976; Porteous, 1990; Healy & Hoyles, 2000). Therefore, there is a gap in research with how students learn to incorporate variables into proof arguments, since high school students avoid them and it is assumed that undergraduate students know how or should know how to use them (e.g. Recio & Godino, 2001). This research suggests that many prospective teachers prefer to use algebraic symbols, but demonstrate multiple limitations in using them appropriately to prove statements. Given the known research from secondary and university students, they may also need more support learning to appropriately use variables when writing proofs.

The second important reason for the low number of proofs involves the development of an argument within a proof. In general, the participants did not struggle to make a generalization, but stating a general argument to support their formula (NxN window: task 4, Sticky Gum: task 6) or solution (Squares: task 2) was a challenge for most them. There were two popular argument types (A2.3 and A3.1), which means that the participants either did not produce a general argument at all (A2.3) or they developed an argument that included assumptions (A3.1). The non-argument solutions were typically generalizations with explanations stating how each part of the formula related to the problem situation, and provided no justification for why the formula works for any situation defined in the problem. The attempts at an argument go beyond explaining the generalization, but the solution lacked a complete argument. Anticipating these shortcomings as the participants develop their solutions or in whole class discussion could support future prospective or in-service teachers with developing complete arguments.

The negotiation between generalization and proof was directly addressed in the course with the sequence of three reasoning-and-proving tasks (G.J. Stylianides & A.J. Stylianides, 2009) in which participants engaged during the first class, and these activities appeared to positively influence the participants understanding that it is insufficient and invalid to base a generalization on a set of examples. For instance, several participants referenced the set of activities and Karen's thinking provides a good representation of what was said:

> *I really like doing this through problems of squares, problems like the 60 by 60, and then the dots on the circle and then the – I mean the counter example is pretty crazy, but I kind of like doing that because I have always been used to finding a pattern, like finding likes – make a conjecture all that, and then we did that circle thing and we saw that it doesn't always hold. I think that's been pretty interesting to just kind of remember that just because it works for a few cases, few situations, it doesn't mean that it's going to hold forever.*

These three tasks intended to promote a cognitive conflict and based on what Karen said it seems as though she now feels as though she needs to make a change. However, the conflict seems to include two levels: 1) the participants come to recognize that a method they previously deemed valid is now insecure; and 2) they have to learn a new acceptable method. Lannin (2005) and others (e.g. Ellis, 2007; Knuth & Sutherland, 2004) have studied the challenges middle school students encounter when asked to provide a valid justification. For example, Lannin explained, "When justifying an algebraic model, an argument is viewed as acceptable when it connects the generalization to a general relation that exists in the problem context" (p. 235). This study shows that some prospective secondary mathematics teachers may also need support learning *how* to connect a generalization to general features of an algebraic model.

The insights into the challenges teachers face in solving proof-related tasks provided by this study can inform the work of teacher educators. Specifically, the reasoning-and-proving codes developed for this research could be used to analyze solutions teachers construct constructed to provide more direct feedback to support them in understanding how to improve their arguments.

The activities used in this study were not useful in detecting growth with critiquing or constructing arguments. The codes were useful in identifying the participants' shortcomings across the various activities. However, after closer inspection, it seems that some of the participants did improve their ability to construct arguments, although this is not evident from examining the codes. For example, Karen became cognizant of the fact that she could rely on her own knowledge of the concepts in the problems instead of trying to follow a particular proof format such mathematical induction. While solving the first task, Karen wrote the following solution:

For any counting number n, $n^2 + n$ *is always even.*
$3^2 + 3 = 9 + 3 = 12$
$4^2 + 4 = 16 + 4 = 20$

$n = 1 \quad 1^2 + 1 = 2$
Assume $n^2 + n$ *even*
$(n + 1)^2 + n + 1$ *will always be even*
$n^2 + 2n + 1 + n + 1 =$
$= \underline{n^2 + n} + 2n + 2$ *will always be even*
even even even

This was the only proof Karen wrote. However, during the interview she explained that she did not understand why this method worked it was just a process she learned in college to produce a proof. When asked to solve the second task (Squares), she also attempted mathematical induction. However, when asked to solve the "O + O = E" task she wrote the following argument:

> *Let n be an even number, so n+1 is an odd number. Suppose you have 2 odd numbers, n+1 and n+3. When you add them together, you obtain (n + 1) + (n + 3) = 2n + 4*
> *Since 2n is divisible by 2, it is even; 4 is also an even number (divisible by 2), so if you add 4 to any even number, the sum is still even (still divisible by 2).*
> *Therefore, the sum of any 2 odd numbers is always even*

While Karen's solution the "O =O = E" task (third problem) is not a proof, it shows that she moved away from a formal method that she did not understand, and applied what she knew about odd and even numbers to make an argument. So even though the codes show regression from A4 to A2, Karen actually made progress in learning that a proof does not need to follow a particular method that she does not even understand.

A second important issue relates to the environments in which participants constructed arguments during this study. The participants were asked to complete the interview tasks individually in 10 to 15 minutes and given about the same amount of time for tasks completed in class. This raises a question about how much time these participants would take to solve a task they planned to implement in their classrooms? In others words, should a participant be labeled as one with a limited ability to construct proofs because she is unable to complete it in 15 minutes? The final question is, how much does a teacher need to be able to do on her own in order to successfully implement a reasoning-and-proving task with her students? If a teacher solved a problem in several ways with a knowledgeable colleague, would this be sufficient to enact the task with students? It might be interesting in future studies of teacher learning to have participants solve some problems outside the classroom and interview environment to understand if added time would improve the number of proofs written.

5.1.3 Critiquing arguments

Overall, the participants were fairly successful in identifying whether or not a solution was a proof and in providing accurate reasons for their choices. However, two important insights emerged: 1) not all 'types' of empirical arguments are convincing; and 2) algebraic arguments that were not proofs were still convincing.

Research suggests that students are convinced by empirical arguments (e.g. Chazan, 1993; Healy & Hoyles, 2000; Porteous, 1990), and some teachers also consider example-based arguments as proof (Healy & Hoyles, 2000; Knuth, 2002a). However, the 'types' of empirical arguments that were convincing to the participants varied. Chazan studied high school students exploring geometric conjectures using a software tool. The empirical arguments in the Healy & Hoyles study were specific sets of numerical examples with no explanation. Knuth had teachers critique two different empirical arguments in which both displayed a single example with a detailed explanation of the particular case. While all of these studies did ask participants to analyze empirical arguments, these examples represent different 'types' of empirical arguments. This current study promoted the inclusion of a generic argument as proof, so distinguishing between empirical arguments where the explanation may seem to be general (Knuth, 2002b) and a generic argument were challenging for some participants. However, no participant identified a set of examples (Healy & Hoyles, 2000) or a generalization without an argument as proof. Therefore, some 'types' of empirical arguments are more challenging than others.

Of the 32 samples of student work analyzed by participants, there was one solution that was not a proof but included correctly manipulated algebraic symbols. However, all but two participants questioned the validity of the argument, and only one participant identified it as non-proof. The more surprising result is that when asked why it is a proof, the participants seemed to employ a different criterion to evaluate the solution. For example, participants indicated that 'they used algebra' or 'algebra is what you want' although it seemed as though they only checked to see that the symbols were manipulated correctly. In other words, they did not seem to try to make sense of the symbols in connection with the problem context. On the other hand, if the argument used everyday language, most participants would read through it several times until they understood what was being said before making a decision. Some teachers claimed that the best arguments were those that were valid and easiest to understand and did not prefer symbols to everyday language for student solutions (Healy & Hoyles, 2000). Other teachers when asked to review an algebraic solution that was not a proof may have "focused on the correctness of the manipulations performed in the argument as opposed to the nature of the argument itself" (Knuth, 2002a, p. 393). So in one instance teachers claimed to favor arguments that made sense to them and in another study teachers make sense of the correctness of the manipulated symbols without attending to how the argument relates to the conjecture. This may be an interpretation issue where "make sense" could just mean no mathematical errors as opposed to trying to understand what the variables mean and to what extent the argument proves the conjecture. It is possible then that most of the participants in this study made sense of the algebra instead of checking to see that the symbolic manipulation provided a valid argument to prove the conjecture.

Given the nature of the instruments used it is not possible to know if the course improved participants' ability to validate solutions. The explanations the participants provided were useful to gain insights into why they identified an argument as a proof or not. However, the student solutions used in the packets were chosen for additional reasons such as to have the participants develop questions to support a particular student solution to advance closer to proof. In a subsequent implementation of the same materials, the student solutions in the interviews were changed to include arguments that the participants in this study produced and included all four argument types (i.e. generic, empirical, rationale, demonstration) to learn if growth can be detected with the participants ability to critique solutions.

5.1.4 Selecting and or modifying reasoning-and-proving tasks

A goal of the course was to prepare prospective secondary mathematics teachers to integrate reasoning-and-proving tasks throughout their curricula so that their future students were provided ongoing opportunities to reason and justify their mathematical knowledge, which are also included expectations in the current standards movement (NCTM, 2000, 2009; CCSSM, 2010). To prepare teachers for this goal, the course design and implementation included a wide variety of activities, in which the most practical for teaching, based on participants comments during the final interview, was the opportunity to select and modify reasoning-and-proving tasks. The rationale for including this activity in the course was that if the participants could select and or modify reasoning-and-proving tasks for instruction throughout their curricula, then students could engage in an integrated curriculum in which proof played a central role and teachers would have a tool for working with any curriculum they encountered.

The overwhelming majority of the selected and modified tasks were classified as having high-level cognitive demand (Stein et al., 2010). The participants in the course that was the focus of this study were in previous graduate classes in which they learned the difference between high and low-level tasks. This course seemed to support teachers in modifying high-level “make a generalization” into a high-level “provide an argument” task. The list of modification principles, which was developed as a whole group during the course was useful in modifying the tasks. An interesting insight is that one modification principle (make observations) was applied at a surface level and another modification ('how do you know') was applied with meaning and purpose.

When analyzing tasks it could be argued that all modifications are surface level, and the implementation of the task is the only way to learn the depth at which a modification is truly understood. Another perspective, and the one that undergirds this research, is that instruction is complicated and many factors can alter a teacher's intention. So another way to understand the extent to which a teacher understands the potential of a question she modified or added to a task is how she answers it herself. The rationale is that a teacher's solution to a task suggests what she considers to be an appropriate answer and it is unlikely that the teacher expects a greater level of sophistication than she herself produced. For example, if a task prompts students to make as many observations as possible, and the teacher solves the task and lists only one observation, then it could be argued that the modification is at a surface level. In other words, any questions added or modified that the teacher does not fully answer could be considered surface level modifications. This particular example occurred in several instances.

The participants were asked to solve their selected and modified tasks in several different ways except for the tasks the participants brought to the third interview. The modification "make as many observations as possible" seemed to be added to tasks at a surface level. Participants that included this modification did not include exhaustive lists of possible observations. On the other hand, the most popular and seemingly surface level modification was to add 'how do you know' as the final question of a task. The teachers in every case answered this question with a proof or non-proof argument. So even though the question may seem to just be added to every task mindlessly, the participants actually expected students to construct an argument. So this leads to the question of why prospective teachers added questions that they did not expect students to fully answer? Two possible rationales are that conventional textbooks are full of questions that have potential the solutions to which typically provide little insight. Another possibility, related to the course is that the participants learned to attach meaning to the question "how do you know." The participants came to understand that this question requires an argument that will be critiqued against a criterion of proof. The course never supported the participants with associating a common meaning to "make observations," and they saw the question as a way to give students access without considering how students would respond. Therefore, with future implementations of the CORP materials more explicit conversations may be needed around why one might include a question in a task and what expectations the questions has for students while prospective or in-service teachers learn to select and modify reasoning-and-proving tasks.

5.2 INTERESTING INSIGHTS THAT EMERGED ACROSS THE RESEARCH QUESTIONS

The implementation of the reasoning-and-proving course changed participants' perceptions of proof and how they thought about teaching it. However, a changed conception of proof does not seem to automatically provide prospective teachers with skills to solve proof tasks, critique arguments (Selden & Selden, 2003), or select appropriate reasoning-and-proving tasks. For instance, Selden and Selden suggest that university students "talk a good line" when asked to explain the process they follow to check whether or not an argument is a proof, but the researchers claim students' talk is a "poor indicator of whether they can actually validate proofs with reasonable reliability" (p. 27). Another interesting comparison is to highlight the difference with how the participants in the current study solved tasks and analyzed student solutions. A strong ability to construct valid arguments seems to positively impact one's ability to critique solutions; however, a limited skill with producing proofs does not necessarily equate to a poor aptitude with judging the validity of an argument. If a teacher possesses a strong ability to develop deductive arguments and dismisses solutions based on the use of diagrams or narrative language, then it would be possible for a teacher to be good at writing proofs and possess a narrow criterion for analyzing them. The next two sections will examine participants' responses and how they completed various activities and compare participants' ability to construct proofs against their skill with critiquing arguments.

5.2.1 Comparing what they said with what they did

This section will compare the analyses from multiple sections in Chapter 4, namely the first section (conceptions of proof), against the analysis reported in research questions three and four. The first part will compare what the participants said about the criteria of proof against how the participants analyzed student solutions. The second section will contrast the purposes for including proof in secondary mathematics that participants reported against the selected and/ or modified tasks. Finally, the third section focuses on the implemented tasks and weighs them against the described equity and opportunity conceptions.

5.2.1.1 Conception of criteria versus critiquing students' solutions

As a group, the participants mentioned all seven characteristics of the criteria of proof throughout the three designated time periods, but some participants identified more characteristics more often. In section three of chapter 4, the results of the participants' ability to distinguish between proof and non-proof arguments were shared. This section will compare the two separate results of what they said versus their skill with critiquing solutions to identify any discrepancies.

The participant (Lucy) who talked the least about the seven criteria of proof characteristics, did not have the lowest ability to critique solutions. The three participants (Karen, Tina, Brittany) with the lowest ability (25 or less correctly identified arguments out of 32) to distinguish between proof and non-proof arguments discussed five or six of the seven characteristics. However, two participants (Katie, Nathaniel) with a high ability (28 or more correct out of 32) to evaluate students' solutions also identified most (6 or 7 out of 7) of the characteristics. So what do these comparisons show? One view could be that talk and ability are unrelated or what Selden and Selden (2003) claim that some students can "talk a good line." Other participants might not be good at articulating their understanding, but are able to apply their knowledge. Another view is that only Katie and Nathaniel really have a full understanding of the criteria of proof for teaching, since they were able to both articulate their understanding and apply it. Therefore, this second perspective acknowledges that both are important and both need to be developed to support teachers with gaining a full criterion of proof.

5.2.1.2 Conception of purpose versus selecting and or modifying tasks

The reason participants were asked about the purpose of proof was to help them begin to think about selecting and modifying reasoning-and-proving tasks to align with their various purposes, and for them to see proof as vital component of mathematics. The participants listed a total of seven purposes throughout the course (shown in table 5.1). Most participants conveyed the goals of the required course tasks as a way to help students learn what is proof (5) and to develop students' ability to develop an argument (7). These two main participant purposes were also goals of the implemented course tasks. While it was the case that the participants were able to identify a broad variety purposes for wanting to enact a proof task, some purposes seem more relevant in practice depending on students' experience with reasoning-and-proving.

Table 5.1. Purposes of proof that the participants identified

Purposes for proof in secondary mathematics
1) To organize definitions and statements
2) To gain a deeper understanding of the truth of mathematics statements
3) To develop logical and rationale thinking skills
4) To learn what is proof
5) To communication mathematical truth
6) To build mathematical authority
7) To develop an ability to construct a proof

During the early portion of a school year, perhaps most of the chosen tasks would be dedicated to supporting students with learning what constitutes a proof. Later in a semester, the teacher would transition into selecting tasks that support students in gaining a deeper understanding of the truth of statements. In other words, students may first need to understand the reasoning skills (find a pattern, make a conjecture, etc,) before this scaffolding is removed in future tasks. This idea surfaces in a task Nathaniel selected in which he wanted students to prove the formula for the area of a triangle. The nature of the task matched purpose two: to gain a deeper understanding of the truth of a mathematical statement. However, he was provided feedback that students may not be able to access the task, since they might not know how to go about developing an argument. He modified the task to include scaffolding so that students would first be required to examine particular cases (reasoning) before developing an argument. The modified version of the task still provided students with an opportunity to gain a deeper understanding of the formula for the area of a triangle, but the added scaffolding included additional purposes for the task that might be more beneficial for secondary students, especially as they gain experience with reasoning.

5.2.1.3 Conception of equity and opportunity versus implementing reasoning-and-proving tasks

All of the participants said they believed all students could engage in reasoning-and-proving opportunities. Six of them believed that teachers should integrate reasoning-and-proving tasks across all topics of all courses, and four participants (Karen, Tanya, Brittany, Katie) communicated an interest in engaging their future students in writing proofs. Two (Karen, Katie) of the four who said they would implement reasoning-and-proving tasks actually did. Attempts to contact Tanya were unsuccessful, but she may have engaged students in solving reasoning-and-proving tasks. Brittany, along with three additional participants (Tina, Lucy, Uma), agreed to the follow-up portion of the research but never returned task packets. A possible reason for the low return rate could be that the new teachers were overwhelmed by their responsibilities at their new schools and struggled to incorporate opportunities into a curricula that may not have included proof tasks. After getting more familiar with the curricula they teach, they may become more comfortable with modifying their curricula to include reasoning-and-proving tasks.

Karen said that proof tasks should be implemented in all secondary courses; explaining that all students were capable and it should be integrated throughout all course topics. However, she did not suggest that she would provide reasoning-and-proving opportunities on a daily basis. Karen was the only geometry teacher at her school and engaged all of her students in reasoning-and-proving tasks. The one discrepancy is that the last two tasks she implemented were pattern tasks, which do not seem to fit a geometry curriculum. Additionally, the geometry tasks she implemented were missing the inductive reasoning quality that she included in her pattern tasks. For example, students were not asked to find a pattern or make a conjecture. Since Karen enacted the pattern tasks (inductive reasoning) after the geometry tasks (deductive reasoning), it seems as though she may not have known how to modify the geometry tasks to include opportunities for students to explore cases.

Katie agreed with Karen's beliefs about students' opportunities to reason-and-prove, stating that proof tasks can be enacted with all students in all course topics. Katie engaged her low-level pre-calculus students in reasoning-and-proving tasks. Since she did not teach geometry, she had a more difficult time identifying tasks that matched the curriculum. She explained that most (3 of 5) of her tasks were implemented between units or before holidays.

Therefore, the challenge for both participants seemed to be finding opportunities to include reasoning-and-proving tasks throughout a secondary curriculum, including geometry. This could mean that the participants believed they were more prepared to include reasoning-and-proving tasks into their curricula than they were or there was a mismatch between what they said and did.

5.2.2 Comparing ability to construct proofs versus critiquing arguments

Selden and Selden asked the question directly, "How does the ability to validate proofs relate to the ability to construct them" (2003, p. 29)? To date, this question has yet to be answered. This current study begins to provide evidence of prospective secondary teachers' abilities and how these two activities are related. The main finding shows that limited skill to validate arguments tends to translate into low ability to construct proofs. The second one is that the converse is not necessarily true; a limited ability to construct proofs does not mean a poor ability to critique arguments. However, the linchpin seems to be the individual's conception of an accurate criterion of proof.

Three participants (Karen, Tina, Brittany) who struggled to construct valid arguments also demonstrated the lowest ability to validate arguments. However, another participant (Katie) who had the greatest ability to validate student solutions did not produce a high number of proofs. The two participants (Nathaniel, Tanya) who wrote the greatest number of proofs also were among the best at critiquing solutions. The difference seems to rest on a participant's conception of proof. Brittany believed that some of her non-proof arguments that she constructed actually counted as proof, which carried over into misidentifying similar student arguments as proof. On the other hand, Katie was well aware that the non-proof arguments she produced were not valid. Other participants who thought the solutions they constructed included assumptions were less critical of assumptions as they read student solutions. Finally, two interesting findings within this comparison are three participants (Karen, Brittany, Uma) who struggled the most with counterexamples (Calling Plans: task 8) did not talk about counterexamples as part of their criteria of proof. Additionally, the two participants (Uma, Tanya) who failed to include a conclusion on three or more arguments when they solved tasks did not mention the need for proofs to have a conclusion. Therefore, not talking about a particular characteristic could mean that the individual is not aware of its importance. Thus, both the feedback participants are provided on the arguments they produce and conversations around validating student solutions contribute to developing a complete criterion of proof. A full understanding of what counts as proof, which supports teachers with knowing what is required, along with additional opportunities to write them, could lead to the construction of more proofs.

5.3 CONCLUSION AND FUTURE RESEARCH IDEAS

A decade ago Knuth (2002b) proposed a challenge:

> Thus, perhaps the greatest challenge facing secondary school mathematics teachers is changing both their conception about the appropriateness of proof for *all* students and their enactment of corresponding proving practices in their classroom instruction (p. 83).

He went on to explain that this is the responsibility of mathematics teacher educators. The current research study was the first to take on both parts of his challenge and was successful at changing prospective teachers' conception of the appropriateness of proof for all students. It is also known that two teachers began to enact tasks related to proof with their students. One shortcoming, which requires additional work, was that the participants did not seem to be fully prepared to integrate reasoning-and-proving tasks seamlessly throughout their curriculum.

The participants explained that the process of selecting and modifying task activities they engaged in during the course was very practical and useful. Most of the participants were successful with selecting high-level "provide and argument" tasks. However, the missing piece seems to be integration. The participants who were not enthusiastic about enacting proof tasks all mentioned time issues; specifically there is already too much other content to cover. In others words, they view proof as an extra topic of study, not essential to learning secondary mathematics. The participants who enacted tasks in their classrooms treated it as a side topic in five out of ten implementations. Therefore, this leads to the question of how might mathematics teacher educators prepare teachers to learn how to integrate proof throughout the secondary curricula? What activities might be most useful?

If the integration of reasoning-and-proof tasks throughout secondary curricula is to occur, then mathematics teacher educators need to first learn what this entails. For instance, is it possible to select any unit out of traditional textbook and identify or modify a task in the chapter to include reasoning-and-proving? If so, then perhaps this should be the focus of task modification. If it is not possible for every unit, then we may need to decide which units are the most appropriate and focus prospective and practicing teacher on that particular content. Aligning tasks with specific units could support teachers with recognizing how it is related to the content they teach. For example, Katie remembered the sequence of three tasks she solved during the course and engaged her own students in a slightly modified version of the tasks when the content (geometric series) surfaced in her curriculum. Ten years later, this course almost met

Knuth's (2002b) challenge. More work is needed, however, to learn the extent to which reasoning-and-proving tasks can be integrated throughout secondary curricula.

In conclusion, recommendations are provided as ways to improve the CORP curriculum materials. In a subsequent implementation of the course, revisions were made to the activities in the interviews. The NxN window task was moved from the second interview (task 4) to the third interview (task 8). The purpose of this was to sequence this problem after participants had solved the Sticky Gum task since both tasks required the solver to make a generalization and explain why the formula is always true. Hence it would be possible to see if work on this ideas in class improved performance on the interview task. The Calling Plans task was removed since the course did not focus on counterexamples, but the results do suggest that prospective teachers need opportunities to learn about counterexamples. The first interview task (N^2 + N is even) was moved to the second interview so that it follows the "O + O = E" task. A Trapezoid Pattern task was used for the first interview. The rationale for using this problem was that it provided a pre-assessment to the Squares problem. The point was to learn if participants would base a generalization on a set of cases or believe that a generalization is proof.

The selected student solutions in the three interviews were altered to better reflect the types of challenges that were encountered in this study. For the two tasks that were retained (N^2 + N is even and NxN window), solutions developed by the participants in this study that aligned with the three argument types (generic, empirical, and rationale) were used. New solutions were designed for the Trapezoid Pattern task that also aligned with the argument type. The intent of these changes was to provide participants with more challenging solutions to analyze and to be able to look across to the arguments types across the interviews to see if any challenges persist or if improvements are made.

Changing some of the student solutions in the course materials may better reflect the three argument types. Many of the student solutions in the Sticky gum set are similar in that they make tables to find a generalization. Since the Sticky Gum problem appears later in the course, it may be useful to include more generic and rationale argument type solutions especially since all of the participants correctly identified more than half (5 of 8) of the current solutions.

Finally, even though most of the participants (8 of 9) were able to select and or modify a task to be high-level "provide and argument," there seems to be a challenge with how to integrate reasoning-and-proving tasks across *all* secondary concepts. It was previously mentioned that this is an important issue to further explore, but a first step in this process might be to create or modify a reasoning-and-proving task for each unit in a secondary curriculum. The next step would be to help teachers apply similar principles to their curricula to create unit reasoning-and-proving tasks.

APPENDIX A

Copies of Each of the Interview Protocols

First Interview

(Interviewer Copy)

Part 1:

Read the following statement so that all teachers are provided the same rationale regarding why they are being interviewed.

Thank you for agreeing to participate in this interview. As you know, we are interested in better understanding your views on reasoning-and-proving and how the course is shaping or reshaping those views. Today's interview has three parts:

Part 1: Respond to a few general questions about proof

Part 2: Create a proof for a mathematical statement

Part 3: Analyze work produced by students when they were asked to create a proof

What I am most interested in is HOW YOU ARE thinking. I will be recording this but we can turn it off anytime.

1.) What experiences have you had with proofs – as a student in high school and college and as a mathematics teacher?

2.) What does it mean to prove a statement?

3.) What should be included in a proof?

4.) What should or could a proof look like?

5.) What role do you think proof should play in the secondary mathematics classroom?

6.) Which courses in the secondary curriculum should or could include work on proofs?

Part 2:

1.) Prove that for every counting number n (1, 2, 3, 4 …), the expression $n^2 + n$ will always be even.[13]

Provide time for interviewee to prove the task. Then ask:

2.) What about your solution makes it a proof?

---

[13] Problem adopted from Morris (2002)

3.) Do you think that there is a counting number n which would cause the expression $n^2 + n$ NOT to be even? Why or why not?

Part 3:

Present the five arguments to the teacher and indicate that these arguments represent students' efforts to create a proof for the task that they themselves have just completed.

Once the participant has had enough time to read through each argument, ask the following question.

Do any or all of the arguments prove that the conclusion is true for each and every counting number? Explain why each of the five arguments is or is not a proof.

Argument 1 - Anne's Solution:

Since $n^2 + n$ can also be written as $n(n + 1)$, then we see that the product represents consecutive numbers. Consecutive counting numbers implies that one of the numbers is even and the other is odd. The product of an odd and even number is even sine one of the numbers is divisible by 2. In other words, n or n+1 divides 2 with no remainder. This implies the product is also divisible by 2. Thus, since $n(n + 1)$ is divisible by 2, it is even. Therefore, $n^2 + n$ is even.

Argument 2 - Ben's Solution:

Let $n = 1$. Then $n^2 + n = 1^2 + 1 = 2$. 2 is even, so this works.

Let $n = 2$. Then $n^2 + n = 2^2 + 2 = 6$. 6 is even, so this works.

Let $n = 3$. Then $n^2 + n = 3^2 + 3 = 12$. 12 is even, so this works.

Let $n = 101$. Then $n^2 + n = 101^2 + 101 = 10{,}201 + 101 = 10{,}302$. 10,302 is even, so this works.

Let $n = 3056$. Then $n^2 + n = 3056^2 + 3056 = 9{,}339{,}136 + 3056$. 9,342,192 is even, so this works.

I randomly selected several different types of numbers. Some were high, and some were low. Some were even and some were odd. Some were prime and some were composite. Since I randomly selected and tested a variety of types of counting numbers, and it worked in every case, I know that it will work for all counting numbers. Therefore, $n^2 + n$ will always be even.

Argument 3 - Cara's Solution:

Let $n = 1$. Then $n^2 + n = 1^2 + 1 = 2$. 2 is even, so it works. Let $n = 2$. Then $n^2 + n = 2^2 + 2 = 6$. 6 is even, so it works. I tried an even and odd number. Since it worked for both an even and an odd number, it will always work. The expression $n^2 + n$ where n is any counting number will always be even.

Argument 4 - Dominique's Solution:

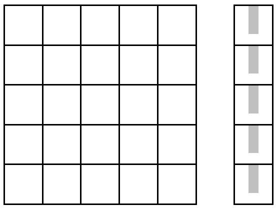
If n is an odd counting number, then n^2 will be odd. An odd plus an odd is even, so since n^2 and n are odd, $n^2 + n$ is even.

If n is an even counting number, then n^2 will be even. An even plus an even number is even, so since n^2 and n are even, $n^2 + n$ is even.

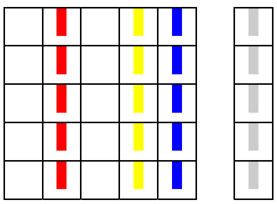
Since all counting numbers are either even or odd, I've taken care of all numbers. Therefore, I've proved that for every counting number n, the expression $n^2 + n$ is always even.

Argument 5 - Edward's Solution:

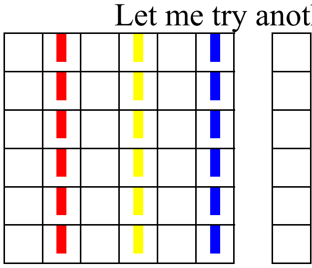
So if I start with a square say 5 by 5 and add it to the number 5



Ok now I will match up the colums so that all but one column has a pair (the blue one). The blue column will be matched with the gray 5 coulmn that is added to the square. So that will make the whole thing even because you can divide the entire thing into two equal pieces.



Let me try another one.



The columns in the 6 by 6 match up perfectly with none left over and the added part 6 folds in half. So every number is paired which makes $6^2 + 6$ an even number.

Now I got it. If the square is an odd by an odd like 5x5, then there will always be a column left over since an odd number does not divide by 2 evenly. The left over column of an odd sided square will always match with the added column part.

If the square is even by even, the every column has a match. The added part for an even by even will also be even based on the problem. And an even number divides two with nothing left over or folds perfectly.

So it does not matter the counting number that you start with when you square it and add it to itself it will always result in an even number.

Part 4:

Did you bring a task with you today?

Why did you select this particular task?

Second Interview

Thank you very much for participating in this second interview. We want to gain insight into your evolving understanding of reasoning-and-proving and of how to help secondary students develop this capacity. The interview will be three parts: general questions about reasoning-and-proving, create a proof, and evaluate student approaches to the same proof task. I will be recording this interview, but I will turn the recorder off at your request any time.

Part 1

1) What do you think is required for an argument to count as proof? Why?

2) How, if at all, has your understanding of reasoning and proving changed over the last four classes?

3) What specific activities do you believe have most helped YOU to better understand reasoning-and-proof?

4) What, if anything, about reasoning-and-proof still is unclear or confusing?

5) How has the course influenced your thinking about teaching reasoning and proving in your classroom?

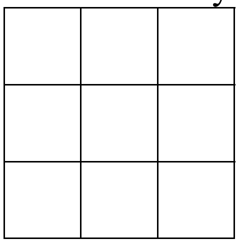
Part 2

The diagram below shows the frame for a window that is 3 feet by 3 feet. The window is made of wood strips that separate the glass panes. Each glass pane is a square that is 1 foot wide and 1 foot tall. Upon counting, you will notice that it takes 24 feet of wood strip to build a frame for a window 3 feet by 3 feet.

1. Determine the total length of wood strip for any size square window.

2. Prove that your generalization works for any size square window.

3ft – by – 3ft Window



Provide time for the interviewee to create a proof. Then ask:

3. What about your solution makes it a proof?
4. Can you think of other possible ways to prove that your generalization works (without writing it out)?

<u>Part 3</u>

Present the four arguments to the teacher and indicate that these arguments represent students' efforts to create a proof for the task that they themselves have just completed.

Ask the participant:

1) Which of the arguments would you classify as proofs? Why?
2) Which argument do you think is most convincing? Why?
3) What questions might you ask Student D to help him in forming a generalization?

<u>Analyze Student Solutions</u>

Student A

Window size x	# of wood pieces y
1 by 1	4
2 by 2	12
3 by 3	24
4 by 4	40
5 by 5	60

I notice that the first difference in the table is 8, 12, 16, and 20 and all of the second differences are 4. Since the second difference is constant (4), then the equation is quadratic.
I know that the y-intercept is 0 since a 0 by 0 window will have zero wood pieces.
Also half of the second difference gives the leading coefficient. Now I just need to find the coefficient for x, which I will call b.

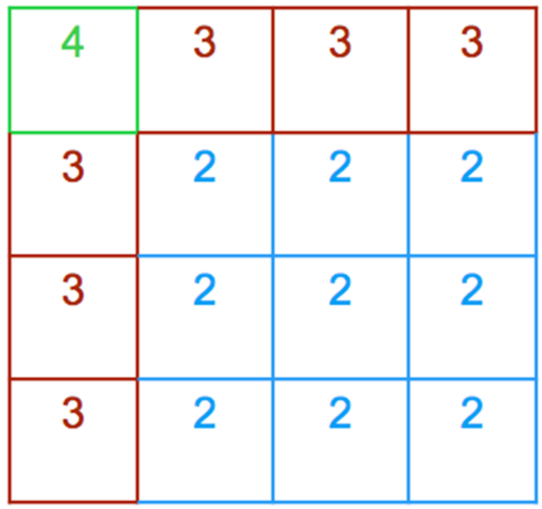
$Y = 2x^2 + bx$

Choosing a random coordinate will allow me to find b. So I will choose (2, 12).
$12 = 2(2)^2 + b(2)$
$12 = 8 + 2b$
$4 = 2b$
$b = 2$

So for any square size window length x, the number of wood pieces is
$2x^2 + 2x$

Student B

I first counted the four wood pieces around the top left windowpane as shown in the diagram. As I move to the right. I noticed that only 3 new wood pieces are being added. I continued this pattern along the top and along the left side. I wrote 2 in the squares that only had two new windowpanes.



From this diagram I know that a 1 x 1 window has 4 wood pieces

A 2 x 2 has 4 + 3 + 3 + 2 = 12

A 3 x 3 has 4 + 3 + 3 + 3 + 3 + 3 + 2 + 2 + 2 + 2 = 24

A 4 x 4 has 4 + 3 + 3 + 3 +3 + 3 + 3 + 3 + 2 + 2 + 2 + 2 + 2 + 2 + 2 + 2 + 2 = 4 + 6(3) + 2 (9) = 4 + 18 + 18 = 40

So there will always be a 4 (in the top left corner) and 2 rectangles of 3s (along the top and along the left side) and a square of 2s.

So for any square there would always be:
1 pane that you counted 4 pieces of wood
(n-1) panes across the top were you counted 3 pieces of wood
(n-1) panes down the side were you counted 3 pieces of wood
$(n-1)^2$ panes were you counted 2 pieces of wood

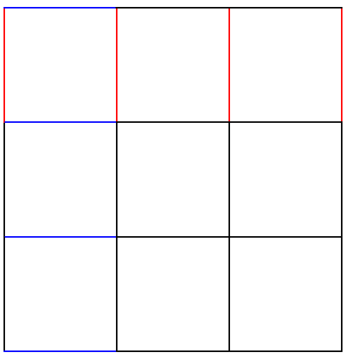
So when you add it all together you get
$4 + 3(n-1) + 3(n-1) + 2(n-1)^2$

Student C

In a 3 x 3 square there are 3 panes across and in each row there are 4 vertical pieces of wood (shown in red). So there is one more vertical piece of wood than there are panes. So the total number of vertical pieces is 12.

There are 3 panes going down and in each column there are 4 horizontal pieces of wood (blue). So there is one more horizontal piece of wood than there are panes. So the total number of horizontal pieces is 12.

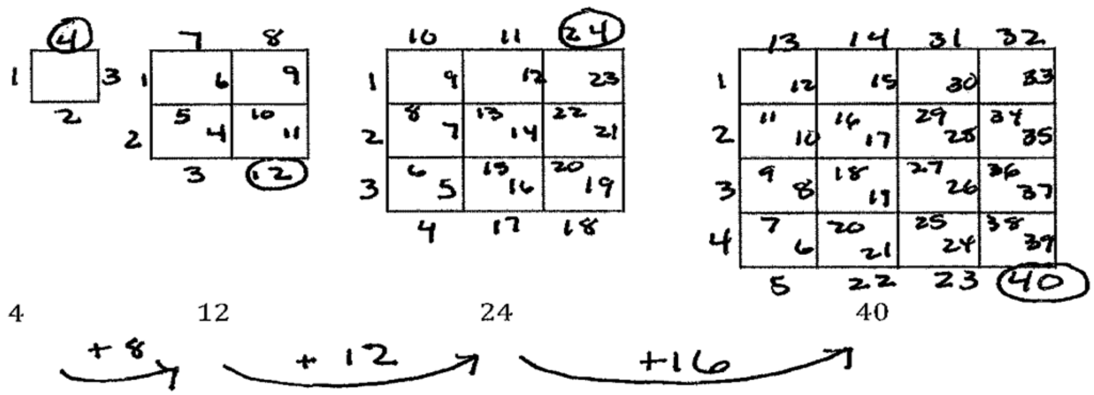
So 12 vertical pieces plus 12 horizontal pieces is 24 pieces and each is a foot long so it is 24 feet total.



So if you have a n x n square, it would have: n panes across and there would be n+1 vertical pieces and n panes down and there would be n+1 horizontal pieces.

So the total number of pieces would be n(n+1) + n(n+1) and this would be the number of feet too because each piece is 1 foot long.

Student D



The first one has 4 and the second has 12 and the third one has 24 and the fourth one has 40. So you add +8, then +12, then +16. And each time you add 4 more than you did the time before. So the fifth one would be +20 and the sixth one would be +24 and so on.

Interview 3

Thank you very much for participating in this third interview. We want to continue to gain insight into your understanding of reasoning-and-proving and your view regarding how to help secondary students develop this capacity. The interview has four parts. In part one you will be asked to discuss what you learned in the course and how you learned it; in part two you will be asked to solve a task and justify your solution; in part three you will be asked to evaluate student arguments; and in part 4 you will be asked to talk about the task you brought with you. I will be recording this interview, but I will turn the recorder off at your request any time.

<u>Part 1</u>

1a. How, if at all, has your understanding of reasoning-and-proving changed over the past six weeks (12 classes)? That is, what is it you understand now that you did not understand prior to taking this class?

1b. What specific activities do you believe have most helped YOU in better understanding reasoning-and-proof? *(Provide teachers with a copy of the course map and ask them to identify specific activities that impacted their learning. For each activity identified, press teachers to explain how the activity caused them to think differently.)*

1c. What, if anything, about reasoning-and-proof still is unclear or confusing?

2a. How has the course influenced your thinking about teaching reasoning-and-proving in your classroom?

2b. What specific activities do you believe have influenced YOUR thinking about teaching students to reason-and-prove? *(Provide teachers with a copy of the course map and ask them to identify specific activities that impacted their thinking about teaching reasoning-and-proving. For each activity identified, press teachers to explain how the activity caused them to think differently.)*

2c. What, if anything, about teaching reasoning-and-proof still is unclear or confusing?

<u>Part 2</u>

1. Provide teachers with a copy of The Calling Plans Task shown below and ask them to answer the questions A and B. *(Note that the task does not ask teachers to create a proof. Part of what we are trying to assess here is whether or not teachers spontaneously produce proofs when asked to explain and justify. They will be asked later if they have produced a proof. If a teacher asks if you want them to create a proof, simply remind the teacher that the task asks them "to explain and justify" and that they should do what ever they think is necessary to satisfy this request.]*

2. Once teachers have completed the task, ask:

   Is your solution a proof? Why or why not?

If not, what would it take to make it a proof?

The Calling Plans Task

Long-distance Company A charges a base rate of $5 per month, plus 4 cents per minute that you are on the phone. Long-distance Company B charges a base rate of only $2 per month, but they charge you 10 cents per minute used.

Keith uses Company A and Rachel uses Company B. Last month, Keith and Rachel were discussing their phone bills and realized that their bills were for the same amount for the same number of minutes. Keith argued that there must be a mistake in one of the bills because they could never be the same. Rachel said that the phone bills could be the same.

C. Who do you think is right, Keith or Rachel? Why?

D. For any two phone plans, is there always a number of minutes that will yield the same cost for both plans? Provide an explanation to justify your position.

Part 3

Provide teachers with copies of student solutions A – E to both questions A and B and ask:

1. Which students do you think provide adequate justification for their position?

2. Which argument do you think is most convincing? Why?

3. Which, if any, of the arguments actually counts as a proof? Why?

4. What questions would you ask Student D to help him make progress on the task?

Student A

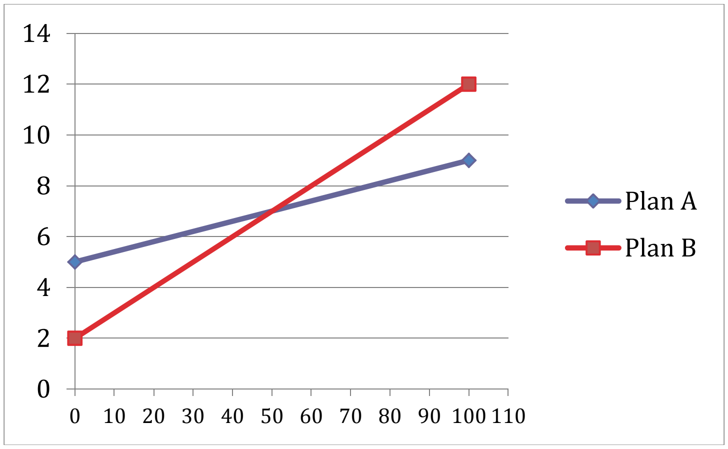
A. I think that Rachel is right because both Company A and B cost $7 for 50 minutes. I figured this out by making a table.

# min	Cost A	Cost B
0	5.00	2.00
10	5.40	3.00
20	5.80	4.00
30	6.20	5.00
40	6.60	6.00
50	7.00	7.00
60	7.40	8.00
70	7.80	9.00
80	8.20	10.00
90	8.60	11.00
100	9.00	12.00

B. If two phone plans don't have the same cost per minute, they will form lines that intersect. If they have the ***same*** *cost per minute they will produce parallel lines that never meet so, NO there is not ALWAYS a number of minutes that gives the same cost.*

Student B

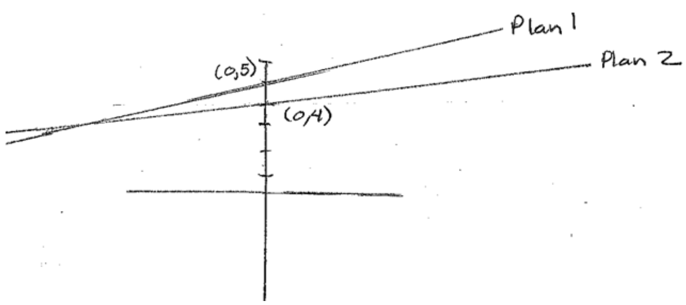
A. I think Rachel is right because her bill and Keith's will be the same when they have talked 50 minutes. I made a graph of both plans and saw that they had a point of intersection at (50, 7).



B. This will not be true for ALL phone plans. If the monthly fee and the cost per minute for Plan 1 are greater than both the monthly fee and the cost per minute for Plan 2, then they will never have the same cost for the same number of minutes.

For example,
Plan 1: cost per minute 4 cents, monthly fee \$5
Plan 2: cost per minute 3 cents, monthly fee \$4

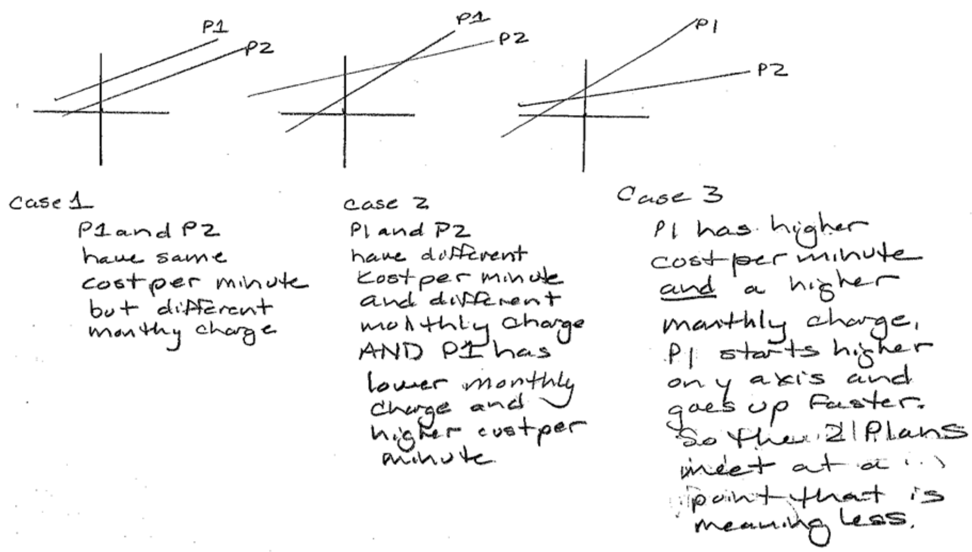
If I graph these two plans they will intersect but not in the first quadrant which is the only one that makes sense when you are talking about phone plans because both the number of minutes and the cost have to have positive values.



Student C
A. Rachel is right. I made two equations and set them equal to each other.
CA = .04m + 5 and CB = .10m + 10
.04m + 5 = .10m + 10
.04m - .04m + 5 = .10m - .04m + 2
5 = .06m + 2
5 – 2 = .06m + 2 – 2
3 = .06m
3 = .06m

.06   .06
50 = m
If I put 50 back in either equation I get 7.
CA= .04 (50) + 5
CA= 2 + 5
CA= 7
So for 50 minutes they are both $7.

B. Two plans will have the same cost for the same number of minutes ONLY when the plan with the lower monthly cost has the higher per minute charge like the picture in case 2. So it is not always true.



Student D

A. I think that Keith is right. If the cost per minute and the monthly fee are different then the plans can't have a value that is the same. A mistake must have been made in figuring out the bills.

B. The number of minutes will never give the same cost for both plans unless the plans have both the same fee and the same cost per minute.

Student E

A. Rachel is right. I used my graphing calculator and put in the two equations
CA = .04m + 5
CB = .10m + 10
And found that the lines intersect at (50, 7) so that means both plans cost $7 for 50 minutes.

B. No. Two plans DO NOT ALWAYS have the same cost for the same minutes. I made two phone plans c_1 and c_2 and set them equal. I found that x (number of minutes) has to be greater than 0 to make sense, so when you subtract the monthly fee and the slopes (cost per minute) you have to have positive values. This ONLY happens when plan 1 has the lower monthly charge and the higher cost per minute.

$c_1 = m_1 x + b_1$

$C_2 = m_2 x + b_2$

$m_1 x + b_1 = m_2 x + b_2$

$m_1 x - m_2 x = b_2 - b_1$

$x(m_1 - m_2) = b_2 - b_1$

$x = \frac{b_2 - b_1}{m_1 - m_2}$

$x \geq 0$ For the 2 phone bills to have a shared cost.

This will only occur when both $b_2 - b_1 \geq 0$ and $m_1 - m_2 > 0$. Therefore, the plan w/ lower monthly cost must have higher per minute charge.

Part 4

Teachers were asked to bring a task that they think would be appropriate for engaging students in some aspect of reasoning-and-proof. Ask them if they brought a task with them today. If they did not bring a task then concluded the interview. If they did bring a task, proceed with the following question IF YOU HAVE TIME. If you do not have enough time, simply collect the task from the teacher and conclude the interview. If you have time, ask the following questions.

1. Why did you select this particular task?

2. With whom would you use this task?

3. How is this task similar to or different from the task you brought to Interview 1 (provide teacher with copy of the task they selected initially)?

4. In what ways did your experiences in the course influence your selection of the task?

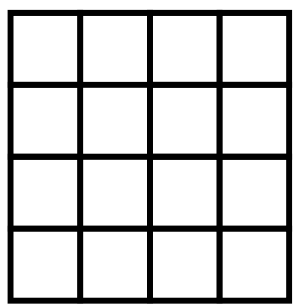
APPENDIX B

Teacher activity sheets for the five course R&P tasks and student work with activity sheets

<u>Sequence of three tasks</u>

The Squares Problem

1. How many different 3-by-3 squares are there in the 4-by-4 square below?



2. How many different 3-by-3 squares are there in a 5-by-5 square?
3. How many different 3-by-3 squares are there in a 60-by-60 square? Are you **sure** that your answer is correct? Why?

The Circle and Spots Problem

Place different numbers of spots around a circle and join each pair of spots by straight lines. Explore a possible relation between the number of spots and the greatest number of non-overlapping regions into which the circle can be divided by this means.

When there are 15 spots around the circle, is there an easy way to tell for sure what is the greatest number of non-overlapping regions into which the circle can be divided?

Looking for a Square Number Problem

Does the expression $1 + 1141n^2$ (where n is a natural number) ever give a square number?

Sum of Two Odds Task

Prove that when you add any 2 odd numbers, your answer is always even.

Analyzing Student Work

Imagine that the students in your class produced responses A-J to the “odd + odd = even” task.

- Review the ten student responses and use the matrix to record whether or not each response qualifies as a proof and provide the rationale that led you to that conclusion.

- Discuss your ratings and rationale with members of your group, come to a group consensus on which responses are and are not proofs and why, and record you group’s decision on the Proof Evaluation Chart.

- Develop a list of criteria for what characteristics an argument must have in order to qualify as a proof.

Student Responses A-J

Student A

If a and b are odd integers, then a and b can be written $a = 2m + 1$ and $b = 2n + 1$, where m and n are other integers.

If $a = 2m + 1$ and $b = 2n + 1$, then $a + b = 2m + 2n + 2$.

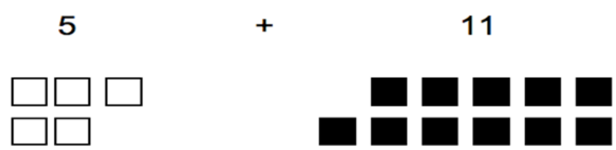
If $a + b = 2m + 2n + 2$, then $a + b = 2(m + n + 1)$.

If $a + b = 2(m + n + 1)$, then $a + b$ is an even integer.

Adapted from: Coxford, A. F., Fey, J. T., Hirsch, C. R., Schoen, H. L., Burrill, G., Hart, E. W., et al. (2003). *Contemporary mathematics in context: A unified approach: Course 3*. New York, NY: Glencoe McGraw-Hill.

Student B

If I take the numbers 5 and 11 and organize the counters as shown, you can see the pattern.

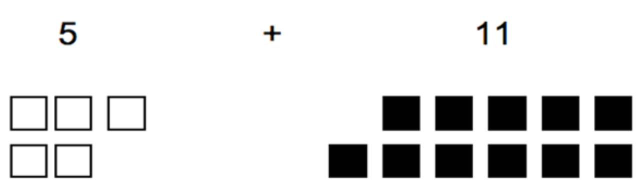


You can see that when you put the sets together (add the numbers), the two extra blocks will form a pair and the answer is always even. This is because any odd number will have an extra block and the two extra blocks for any set of two odd numbers will always form a pair.

Adapted from: Coxford, A. F., Fey, J. T., Hirsch, C. R., Schoen, H. L., Burrill, G., Hart, E. W., et al. (2003). *Contemporary mathematics in context: A unified approach: Course 3*. New York, NY: Glencoe McGraw-Hill

Student C

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.

Adapted from: Coxford, A. F., Fey, J. T., Hirsch, C. R., Schoen, H. L., Burrill, G., Hart, E. W., et al. (2003). *Contemporary mathematics in context: A unified approach: Course 3.* New York, NY: Glencoe McGraw-Hill.

Student D

An odd number = [an] even number + 1. e.g. 9 = 8 + 1
So when you add two odd numbers you are adding an even no. + an even no. + 1 + 1. So you get an even number. This is because it has already been proved that an even number + an even number = an even number.
Therefore as an odd number = an even number + 1, if you add two of them together, you get an even number + 2, which is still an even number.

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

Student E

Any odd number can be written as 2x + 1. So let's add two odd numbers.

2x + 1 + 2x + 1 = 4x + 2
4x + 2 is even since 4 and 2 are both even.
Or 2(2x + 1) shows that 4x + 2 is even.

Student F

$3a + 3b = 6(a + b)$ for $a = 3$; $b = 9$

(3 * 3) + (3 * 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.

Student G

An odd number has to have an odd digit in the ones place. When you add any two single digit odd numbers you would get an even number in the ones place. So here are all of the numbers you get when you add two single digit odd numbers.

	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

The ones place is the only place that matters in determining if a number is odd so it doesn't matter how many other digits it has. If it is odd it will always have a 1, 3, 5, 7, or 9 in the ones place

Student H

My answer

add 1 (a)	add 2 (b)	a + b I
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.

Student I

If you add two odd numbers, the two ones left over from the two odd numbers (after circling them by twos) will group together to make an even number.

Adapted from: Coxford, A. F., Fey, J. T., Hirsch, C. R., Schoen, H. L., Burrill, G., Hart, E. W., et al. (2003). *Contemporary mathematics in context: A unified approach: Course 3*. New York, NY: Glencoe McGraw-Hill.

Student J

Definition of an even number: An integer *p* is even if and only if there is an integer *k* such that *p* = 2*k*

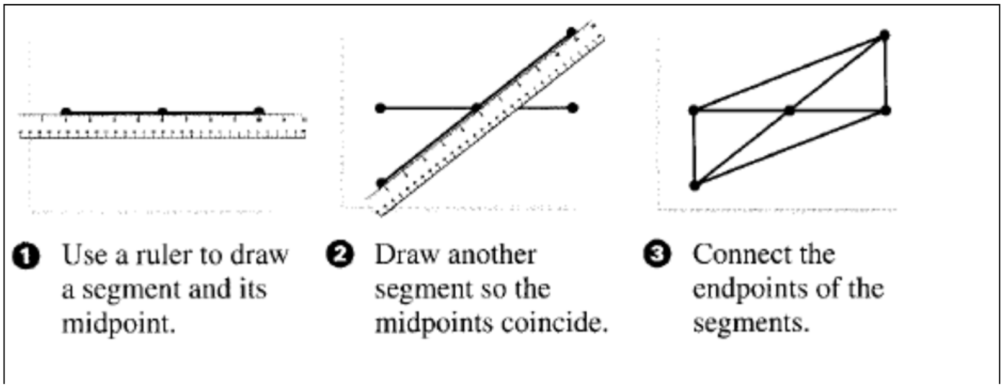
Definition of an odd number: An integer *q* is odd if an only if there is an integer *k* such that *q* = 2*k* + 1. Let's assume X and Y are odd where X = 2*n* + 1 and Y = 2*m* + 1, and n and m are integers.

Statement	Reason
X is an odd number	Given
X = 2*n* + 1	Definition of odd number
Y is an odd number	Given
Y = 2*m* + 1	Definition of an odd number
X + Y = 2*n* + 1 + 2*m* + 1	Addition Property of Equality
X + Y = 2*n* + 2*m* + 1 + 1	Commutative Property of Addition
X + Y = 2*n* + 2*m* + 2	Substitution
X + Y = 2 (*n* + *m* + 1)	Distributive Property
n + *m* + 1 is an integer	Closure Property of Addition for Integers
X + Y is an even number	Definition of an even number

Activity Sheet 4.2

Construction Conjectures

Consider the construction below.



(Adapted from McDougal Littell (2004), *Geometry*, p. 343, #29)

Record your work on the following questions in your notebook or binder.

A. Use this construction with a variety of starting segments. What type of figure does the construction produce?

B. Using the results, make a mathematical argument that explains why that figure is produced each time by the construction.

C. Create a new construction that also begins with a segment and its midpoint but is different in some way. What generalization can you make about any figure created by this construction?

Activity Sheet 5.1

Solving a Mathematical Task: 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 in 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.

Why is three cents the most she will have to spend to satisfy her twins?

1) 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?

2) 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?

3) 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.

Activity Sheet 5.2

Evaluating Student Responses

Imagine that the students in your class produced responses A-H to *A Sticky Gum Problem*.

- Review the eight student responses and determine which of the students actually produced a proof. (Use the *Criteria for Judging the Validity of Proof* from Activity 2.2 to justify your selections.)

- Discuss your ratings and rationale with members of your group, come to a group consensus on which responses are proofs and why, and record you group's decision on the Proof Evaluation Chart.

- As a group, select one response that you think is "close" to being a proof and determine what is missing and what questions you could ask to help the student make progress.

A

(4) Many generalizations can come from these problems

children	colors	money
2	2	3
2	3	4
2	4	5
2	5	6
2	6	7

children	colors	money
3	2	5
3	3	7
3	4	9
3	5	11
3	6	13

children	colors	money
4	2	7
4	3	10
4	4	13
4	5	16
4	6	19
x	y	n

n = xy − (previous y)

Multiply the number of children by the number of colors and subtract the previous amount of colors.

ex

child	colors	money
4	2	7
4	3	10
4	4	13
4	5	16

4x4−3 = 13

B

For problem number four, I drew in and out machines to see what kind of equations I could come up with.

For two children:

# of candies	money
2	3
3	4
4	5
5	6

For three children:

# of candies	money
2	5
3	7
4	9
5	11

Solution: 2. Ms. Hernandez would have to spend fouc cents on candy for her children.

3. Mr. Hodges would have to pay seven cents for candy for his children.

4. It can be said that to find out how much money you would have to pay with three two children is just the number of the different colors of candy plus one (x+1, where x= the number of different colors of candy). To find out how much you would have to pay for three children, it is the number of different kinds of candy multiplied by two and added to one (2x+1 , where x= the number of different colors of candy).

C

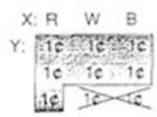
Here is the formula needed to rewrite problem 4 algebraically:

x = colors
y = children
z = cents

$$xy - (x - 1) = z =$$

$$3 \cdot 3 - (3 - 1) =$$
$$9 - 2 = 7¢$$

The reason I chose this formula is as follows: I needed to multiply the colors by the children in order to get the maximum amount of money needed (including children getting more than one color of the same color). But since the children only have to have the same color as one of the gum balls, I needed to take away the other two possibilities, which is why I subtracted the colors minus 1. Look at the following diagram on the next page:



See, we don't need the last two results, of the triplets getting the same color of all the gum balls, just one color -- which is why we subtract the last two numbers, by taking the number of colors, and subtracting one, which in this case is 3 - 1, giving us 2, which we subtract from the kids times the colors, resulting in $3 \cdot 3 - (3 - 1) = 9 - 2 = 7$. Whew! Long sentence!

D

Now Mr. Hodges with his 3 kids comes along to the 3 color machine well I'm getting sick of writing out ~~[illegible]~~ possibilities I'm going to start on the Real POW now!

O.K. says my mind time for an in, out machine

K	C	a
2	2	3
2	3	4
3	3.	7

turn over ↑

K = # of kids C = # of colors a = how much $ ¢ you have to spend.

after making this in, out machine I thought WOW ~~[illegible]~~ ~~[illegible]~~ each answer is 1 less than the K and C added. that died as soon as I did Mr. Hodges ~~[illegible]~~ triplets but his was only one above so I spazzed out on the original in, out machine I did and discovered that all the answers were either 1 above or 1 less than K+C so after brainstorming for awhile.

this is what I came up with for a final answer

$(K-1)C+1=A$ I then tried it with all the #'s I had in my in, out machine and it worked Yaaah!

E

4) Example for any amount of gumballs to match the color and amount

R- red
B- blue
W- white

Z- any color (r b or w)

1¢	1¢	1¢	1¢	1¢	1¢	1¢	1¢	1¢	1¢
r	b	w	r	b	w	r	b	w	Z
1	1	1	2	2	2	3	3	3	4

a b c d

letters stand for the names of quadruplites

My generalization

I think that no matter how many of the same color you need you can always find out what the maxime cost will be. You just figure out how many gumballs with the same color is needed. For example if the amount you need is ten the same.

You need to figure out how many different colors are inside the machine. You take the number you need (in this case ten) and multiply it by the number below the number you need (it would be nine) The answer is 90. Then it doesn't matter what color comes out it will match to make a ten. All it takes is one more making it 91.

Example of work:

number needed - 10
one less - x 9
90) multiply

any color from machine - 90
+ 1
91) add one to the 90

F

4. 3 colors of gum

# of kids	1	2	3	4	5	n	0th term
¢	1¢	4¢	7¢	10¢	13¢	-2+3n	-2

2 colors of gum

# of kids	1	2	3	4	5	n	0th term
¢	1¢	3¢	5¢	8¢	11¢	-1+2n	-1

Formula → x = # of kids y = # of different colors

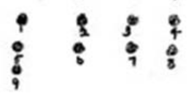
$(1-y)+yx$

*2 Process: I solved the problem by making two in-out tebles. These tables gave me different chunks of infermation to put together. The first was used to find the relationship between the number of kids money. The second was used to find the relationships between the kids, money, and number of different colors of gum balls. The hardes part was starting, beccuse I didn't knaw exactly

G

For parents with more than two children, we will use a visual representation. In order to require a maximum number of gumballs, the machine must first produce one gumball of each color, then a second of each color, a third of each color, and so on. If there are 4 colors in the machine, and 3 children, the max # of gum balls can be represented as:

color 1 2 3 4



each row gets completely filled before a new row is started. As soon as there is one gum ball in the final row, the third row for the third child, there will be same color of gum ball with 3 of that color, one for each child. Each row before the row for the final child must be completely filled, meaning (the number of children − 1) times the number of colors must be bought, + 1 final gumball for the last child. The formula is ∴ $\max\$ = (k-1)g + 1$

where k = # of kids

and g = # of colors

H

In order to generalize,
group the in sets of colors
so we have a machine with
three colors, red, green and
blue, in a worst case scenario,
you will draw R, G, B or
some different permutation
of those colors. If you have
three kids, you will keep drawing
and again receive o permutation
of R, G, B. Then the next
one you draw, no matter what
color, will catch up with one
set of two colors you previously
drew, giving you three gumballs of
the same color. So, you have
drawn out n-1 sets of gumball
colors, and you draw 1 more
giving you $(n-1)g + 1$.

Activity Sheet 6.1A

Explaining Number Patterns

1. Pick any two consecutive whole numbers
2. Square each number and subtract the smaller sum from the larger
3. Add the two original numbers together
4. Make a conjecture about the numbers you found in #2 and #3 (try more examples if you like!)
5. Prove that the conjecture you made in #4 will always be true. *After you have proven the conjecture in one way, see if you can prove it using another strategy or method.*

Adapted from Slavit, D. (2001). Revisiting a difference of squares. *Mathematics Teaching in the Middle School*, *6*(6), 378-381.

APPENDIX C

Reasoning-and-Proving Task Collection Packet

Please collect artifact packets on reasoning-and-proving lessons.

Each artifact packet includes:

1) The original task as you found it and if you modify it the modified version as well. *(1 copy of each)*

2) See Sorting Students Solution Sheet

3) Any thing that you created in preparing or in enacting the lesson (e.g. lesson plan, solutions, lists your class created, power point, etc.)

4) R&P Task Cover Sheet (*complete for each task*) (Attached Next Page)

5) Background sheet (*complete once*) (Attached after R&P Task Cover Sheet)

Name ________________________________ Task Name ___________________

R&P Task Cover Sheet

Number of Students in Class______

Date Implemented __________

Please use the space provided as a general guide for the length of your answers. If you need additional space please use the back of this sheet.

1. What mathematics unit was the class studying when this task was implemented? Where did you find the task (textbook or another resource)?

If you made modifications, please explain your rationale.

<u>Reflection</u>

Describe any directions, oral or written, you gave to students that are not included on the task itself. Please explain any expectations you relayed to your class. (e.g. Did students work in groups?, Did you grade their work?)

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 enactment?

Explain your overall reaction to your implementation with this task. (What do you believe the students' learned or you learned, would you teach this task again, etc.

Background Sheet

The responses to these questions are meant to be your current perceptions and not questions that you need to investigate. So if you are unsure, then please indicate it.

1. For the class you are collecting data:

Number of students ______

Title of class & grade level (please indicate if it is honors, remedial, etc.)

Did you grade any of the proof tasks? If so explain your method (attach a rubric if you created one)

Did the class develop a criterion for judging proof? If so, please attach.

2. Describe the extent and ways in which colleagues in your department engage their students in R&P and/or support your efforts.

4. Describe the extent and ways in which your school or district supports R&P activities.

BIBLIOGRAPHY

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

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.

Bell, A. (1976). A study of pupils' proof – explanations in mathematical situations. *Educational Studies in Mathematics*, 7, 23-40.

Bieda (2010). Enacting proof-related tasks in middle school mathematics: challenges and opportunities. *Journal for Research in Mathematics Education,* 41(4), 351-382.

Boaler, J., & Humphries, C. (2005). *Connecting mathematical ideas: Middle school video cases to support teaching and learning.* Portsmouth, NH: Heinemann.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*. 33(8), 3-15.

Borko, H., & Putnam, R.T. (1995). Expanding a teacher's knowledge base: A cognitive psychological perspective on professional development. In T.R. Guskey & M. Huberman (Eds.), *Professional development in education: New paradigms & practices* (pp. 35-66). New York: Teachers College, Columbia University.

Boston M.D. (2006). *Developing Secondary mathematics teachers' knowledge of and capacity to implement instructional tasks with high-level cognitive demands.* Unpublished doctoral dissertation, (University of Pittsburgh). UMI Dissertation Services, #3223943.

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, 119-156.

Boston M.D. & Smith M.S., (2011). A 'task-centric approach' to professional development:

enhancing and sustaining mathematics teachers' ability to implement cognitively challenging mathematics tasks. *ZDM Mathematics Educations*, 43, 965-977.

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

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.

Chazan, D. (1993). High school geometry students' justification for their views of empirical evidence and mathematical proof. *Educational Studies in Mathematics*, 24, 359-387.

Chazan, D. (1990). Quasi-empirical views of mathematics and mathematics teaching. *Interchange*, 21(1), 14-23.

Coe, R. & Ruthven, K. (1994). Proof practices and constructs of advanced mathematics students. *British Educational Research Journal.* 20(1), 41-53.

Cohen D.K. (1990). A revolution in one classroom: The case of Mrs. Oublier. *Educational Evaluation and Policy Analysis,* 12(3), 311-329.

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

Darling-Hammond, L., Wei, R. C., & Orphanos, S. (2009). *Professional learning in the learning profession: A status report on teacher development in the United States and abroad*: National Staff Development Council.

Doyle, W. (1988). Work in mathematics classes: The context of students' thinking during instruction. *Educational Psychologist, 23*(2), 167-180.

Doyle, W. (1983). Academic work. *Review of Educational Research, 53*(2), 159-199.

Edwards, B.S., & Ward, M.B. (2004). Surprises from mathematics education research: Student (mis)use of mathematics definitions. *The American Mathematical Monthly,* 111(5), 411 -424.

Fennema, E., Carpenter, T., Franke, M., Levi, L., Jacobs, V., & Empson, S. (1996). A longitudinal study of learning to use children's thinking in mathematics instruction. *Journal for Research in Mathematics Education, 27*(4), 403-434.

Firestone W.A., Mangin M.M., Martinez M.C., & Polovsky T., (2005). Leading coherent professional development: A comparison of three districts. *Educational Administration Quarterly, 41(3), 413-448.*

Greeno, J.G. (2006). Theoretical and practical advances through research on learning. In J.L.

Green, G. Camilli, & P.B. Elmore (Eds.), *Handbook of complementary methods in education research,* (pp. 795 – 822). Washington, DC, and Mahwah, NJ: American Educational Research Association and Lawrence Erlbaum.

Greeno, J.G. (1997). On claims that answer the wrong questions. *Educational Researcher,* 26(1), 5-17.

Greeno, J.G. (1993). Comments on Susanna Epp's chapter. In A. Schoenfield (Ed.), *Mathematical Thinking and Problem Solving*, (pp. 282-290). Hillsdale, NJ: Lawrence Erlbaum Associates.

Guskey T.R., & Yoon K.S., (2009). What works in professional development? *Phi Delta Kappan*, 90(7), 495-500.

Hanna, G. (1995). Challenges to the importance of proof. *For the learning of mathematics,* 15(3), 42-49.

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

Hanna, G. (1983). *Rigorous proof in mathematics education.* Toronto, Ontario: OISE Press.

Hanna, G. and Jahnke, H.N. (1996). Proof and proving. In A. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook of mathematics education, 877 - 908.* Dordrecht: Kluwer.

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. (2007). Toward comprehensive perspectives on the learning and teaching of proof. In F. Lester (Eds.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 805-840). Greenwich, CT: Information Age.

Harel, G. & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In A. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), Research in collegiate mathematics education, III (pp. 234-283). Washington DC: Mathematical Association of America.

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

Henningsen, M. & Stein, M. K. (1997). Mathematical tasks and student cognition: classroom based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education*, 28 (5), 524-549.

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

Hiebert, J., & Grouws, D. (2007). *The effects of classroom mathematics teaching on student* *learning.* In F.K. Lester (Ed.), Second handbook of research on mathematics teaching and learning. Greenwich, CT: Information Age Publishing.

Housman D. & Porter M. (2003). Proof schemes and learning strategies of above-average mathematics students. *Educational Studies in Mathematics,* 53(2), 139–158.

Johnson, G. J., Thompson, D. R., & Senk, S. L. (2010). A framework for investigating proof related reasoning in high school mathematics textbooks. *Mathematics Teacher, 103*, 410 -418.

Jones, K. (1997). Student-teachers' conceptions of mathematical proof. *Mathematics Education Review,* 9, 21-32.

Jones, K. (2000), The student experience of mathematical proof at university level. *International Journal of Mathematical Education*, **31**(1), 53-60.

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

Knuth, E. (2002b). Teachers conceptions of proof in the context of secondary school mathematics. *Journal for Research in Mathematics Education,* 5(1), 61-88.

Knuth, E. & Sutherland, J. (October, 2004). Student understanding of generality. *Proceedings of the Twenty-sixth Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*, 561-567.

Lannin, J.K. (2005). Generalization and justification: The challenge of introducing algebraic reasoning through patterning. *Mathematical Thinking and Learning*, 7(3), 231-258.

Lappan, G., Fey, J. T., Fitzgerald, W. M., Friel, S., & Phillips, E. D. (2006). *Connected Mathematics 2.* Boston: Pearson Prentice Hall

Martin, G.W. & 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 students actions in the teaching and learning of geometric proof. *Educational Studies in Mathematics, 60,* 95-124.

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

Morris, A.K. (2002). Mathematical reasoning: Adults' ability to make the inductive-deductive distinction. *Cognition and Instruction,* 20(1), 79-118.

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

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

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

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

National Council of Teachers of Mathematics (1980). *An Agenda for Action.* Reston, VA: NCTM.

National Governors Association Center for Best Practices, Council of Chief State School Officers (2010). Common Core State Standards for Mathematics. Washington, DC: Author.

Porteous, K. (1990). What do children really believe? *Educational Studies in Mathematics,* 21, 589-598.

Pólya, G. (1988). First Princeton Science Library Edition: Forward by John H. Conway. How to Solve It: A New Aspect of Mathematical Method. United States: Princeton University Press.

Recio A.M., & Godino J.D. (2001). Institutional and personal meaning of mathematical proof. *Educational Studies in Mathematics,* 48, 83-99.

Reid, D. (2005). The meaning of proof in mathematics education. Paper presented to Working Group 4: Argumentation and Proof, at the Fourth annual conference of the European Society for Research in Mathematics Education. Sant Feliu de Guíxols, Spain. 17 – 21 February 2005.

Scher, L., & O'Reilly, F. (2009). Professional development for K-12 math and science teachers: What do we really know? *Journal of Research on Educational Effectiveness, 2*(3), 209 - 249.

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.

Senk, S. (1985). How well do students write geometry proofs? *Mathematics Teacher*, 78(6), 448 456.

Shulman, L.S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1-22.

Silver, E. A. & Cai, J. (1996). An analysis of arithmetic problem posing by middle school students. Journal for Research in Mathematics Education, 27(5), 521-539.

Simon, M.A. & Shifter, D. (1991). Towards a constructivist perspective: An intervention study of mathematics teacher development. *Educational Studies in Mathematics, 22*(4), 309-331.

Smith, J. C. (2006) A sense making approach to proof: Strategies of students in traditional and problem-based number theory courses. *Journal of Mathematical Behavior,* 25, 73-90.

Smith, M.S. & Stein, M.K. (2011). Five practices for orchestrating productive mathematics discussions. Reston, VA, and Thousand Oaks, CA: National Council of Teachers of Mathematics, and Corwin.

Smith, M.S. (2000). Balancing the old and new: An experienced middle school teacher's learning in the context of mathematics instructional reform. *The Elementary School Journal, 100*(4), 351-375.

Smith, M.S., & Stein, M.K. (1998). Selecting and creating mathematical tasks: From research to practice*. Mathematics Teaching in the Middle School, 3*(5), 344-350.

Solomon, Y. (2006). Deficit or difference? The role of students' epistemologies of mathematics in their interactions with proof. *Educational Studies in Mathematics, 61*, 373-393.

Sowder, L. & Harel, G. (1998). Types of students' justification. *The Mathematics Teacher,* 91, 670-675.

Spillane, J.P. & Jennings, N.E. (1997). Aligned instructional policy and ambitious pedagogy: Exploring instructional reform from the classroom perspective. *Teachers College Record*, 98(3), 449-81.

Spillane, J. & Zeuli, J. (1999). Reform and teaching: Exploring patterns of practice in the context of national and state mathematics reforms. *Educational Evaluation and Policy Analysis,* 21(1), 1-27.

Stein, M. K., Grover, B., & 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., Smith, M.S., Henningsen, M., & Silver, E.A. (2000). Implementing standards-based mathematics instruction: A casebook for professional development. New York: Teacher College Press.

Stigler, J. W., & Hiebert, J. (2004). Improving mathematics teaching. *Educational Leadership, 61(5)*, 12-17.

Stylianides, A. J. (2009). Breaking the equation "empirical argument = proof." *Mathematics Teaching,* 213, 9-14.

Stylianides, A.J. (2007). Proof and proving in school mathematics. *Journal for Research in Mathematics Education,* 38, 289-321.

Stylianides, G. J. (2010). Engaging secondary students in reasoning-and-proving. *Mathematics Teaching, 219* (September), 39-44.

Stylianides, G.J. (2009). Reasoning-and-proving in school mathematics textbooks. *Mathematical Thinking and Learning,* 11, 258-288.

Stylianides, G. J. (2008). An analytic framework of reasoning-and-proving. *For the Learning of Mathematics,* 28(1), 9-16.

Stylianides, G. J., & Stylianides, A. J. (2009). Facilitating the transition from empirical arguments to proof. *Journal for Research in Mathematics Education*, 40, 314-352.

Stylianides, G. J., Stylianides, A. J., & Philippou, G. N. (2007). Preservice teachers' knowledge of proof by mathematical induction. *Journal of Mathematics Teacher Education*, 10, 145 -166.

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

Thompson, C.L., & Zeuli, J.S. (1999). The frame and 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.

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

Weber, K. (2008). How mathematicians determine if an argument is a valid proof. *Journal for Research in Mathematics Education*, 39, 431-459.

Wei, R. C., Darling-Hammond, L., and Adamson, F. (2010). Professional development in the United States: Trends and challenges. Dallas, TX. National Staff Development Council.