INVESTIGATING VIABLE ARGUMENTS: PRE-SERVICE MATHEMATICS TEACHERS' CONSTRUCTION AND EVALUATION OF ARGUMENTS

by

Kim Nordby

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Mathematics

MONTANA STATE UNIVERSITY
Bozeman, Montana

July 2013

©COPYRIGHT

by

Kim Nordby

2013

All Rights Reserved

APPROVAL

of a dissertation submitted by

Kim Nordby

This dissertation has been read by each member of the dissertation committee and has been found to be satisfactory regarding content, English usage, format, citation, bibliographic style, and consistency and is ready for submission to The Graduate School.

Dr. Jennifer Luebeck

Approved for the Department of Mathematical Sciences

Dr. Kenneth Bowers

Approved for The Graduate School

Dr. Ronald W. Larsen

STATEMENT OF PERMISSION TO USE

In presenting this dissertation in partial fulfillment of the requirements for a doctoral degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library. I further agree that copying of this dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for extensive copying or reproduction of this dissertation should be referred to ProQuest Information and Learning, 300 North Zeeb Road, Ann Arbor, Michigan 48106, to whom I have granted "the right to reproduce and distribute my dissertation in and from microform along with the non-exclusive right to reproduce and distribute my abstract in any format in whole or in part."

Kim Nordby

July 2013

DEDICATION

This dissertation is dedicated to Christina, Reece, Anna, Marina, and Auguste.

ACKNOWLEDGEMENTS

I would like to express my gratitude to the members of my graduate committee for their accessibility and insightful comments. Special thanks goes to my chair, Dr. Jennifer Luebeck, for the time, support, guidance, and advice she has offered throughout this process. Thanks also to Dr. Elizabeth Burroughs for her insightful comments and guidance, Dr. Brian Lindaman and Dr. Warren Esty for their willingness to assist and thoughtful advice, and to Dr. John Borkowski for his support and guidance. I would also like to express my thanks to Dr. Maurice Burke for all the encouragement he has provided throughout the years and to Dr. David Yopp for helping lay the foundation for this research project.

Special thanks also go to the five pre-service teachers that agreed to participate in the research. Without their willingness to participate there would be no data.

Last but not least I would like to thank my family. My parents for their support and belief in me. My fiancée Christina for her love and support, and our children, Reece, Anna, Marina, and Auguste for being the greatest kids. I cannot describe how much it has meant to me. I love you.

TABLE OF CONTENTS

1. STATEMENT OF PROBLEM....1

Introduction....1
Background....5
Problem Statement....8
Purpose of the Study....9
Research Questions....10
Significance of the Study....11
Limitations of Study....13

2. REVIEW OF LITERATURE....15

Introduction....15
The Language of Proof, Reasoning, and Argumentation....16
Definition of Terms....16
Vocabulary....18
An Overview of Proof....21
Proof Categories....21
Levels of Proof and Types of Argument....22
Proof Schemes....26
The Empirical Deductive Distinction....27
The Role of Examples....30
Argumentation in the K-12 Classroom....31
Roles and Purposes of Proof....31
Definitions of Proof for School Mathematics....32
Problems with Definitions of Proof for School Mathematics....35
Proof versus Viable Argument....36
Teaching Proof, Reasoning, and Argumentation....37
Teachers' Understanding of Proof and Argumentation....37
Recommendations for Teacher Education....40
Conclusion....41

3. METHODOLOGY....43

Introduction....43
Research Questions....43
Research Design....44
The Researcher....47
Participants and Mathematical Coursework....48
Participants....49
Methods of Proof....50

TABLE OF CONTENTS – CONTINUED

Higher Mathematics for Secondary Teachers....51
Modern Geometry....51
Sample....52
Instrument Development....54
Selection of Construction Tasks....55
Considerations of the Population....55
Consideration of Difficulty Level....56
Consideration of Reliability....58
Selection of Evaluation Tasks....59
Description of Instrument and Tasks....64
Uncovering Pre-Service Teachers' Conceptions of Argumentation....64
Interview Items for Construction....65
Interview Items for Evaluation....68
Pilot Study....71
Results from the Pilot Study....73
Interpretation Tasks....73
Evaluation Tasks....74
Construction Tasks....75
Inconsistencies....76
Data Analysis....76
How the Pilot Study informs Data Analysis....77
Initial Analysis....78
Analysis of Responses to Interpretation Tasks....79
Analysis of Responses to Construction Tasks....81
Analysis of Responses to Evaluation Tasks....84
Issues of Validity and Reliability....85
Conclusion....88

4. RESULTS....89

Purpose of the Study....90
Rationale for Data Collection, Analysis, and Reporting....91
Review of Participants and Program....93
Individual Profiles....94
David....94
Bob....96
Bridget....98
Sam....99
Linda....101

TABLE OF CONTENTS – CONTINUED

Analysis of the Initial Task 102
David 102
Bob 104
Bridget 106
Sam 108
Linda 109
Analysis across Cases 111
Analysis of Viable Argument Items 112
David 113
Bob 116
Bridget 120
Sam 123
Linda 125
Analysis across Cases 127
Viable Arguments 127
Relationship to Proof 129
Role of Viable Argument in School Mathematics 129
Analysis of Construction Items 130
David 130
Bob 133
Bridget 135
Sam 137
Linda 140
Analysis across Cases 142
Analysis of Evaluation Items 144
Faculty Member Analysis 144
David 147
Bob 151
Bridget 154
Sam 157
Linda 162
Analysis across Cases 165
Conclusion 166

5. CONCLUSIONS 168

Introduction 168
Overall Research Conclusions 168
Defining and Constructing 169
Viable or Not? 169
Constructing vs Evaluating Arguments 171

TABLE OF CONTENTS – CONTINUED

Viable Argument vs Proof 172
Empirical vs Deductive Justifications....173
Conclusions Based on Research Questions 174
Conclusions for Research Question 1 174
Research Question 1 174
Conclusions for Research Question 2 177
Research Question 2 177
Logical Progressions....177
Examples....178
Conclusions for Research Question 3 179
Research Question 3 179
Flawed Logic 180
Improvements 181
Examples....181
Implications for Teacher Education....182
Recommendations for Future Research 184
The Influence of Coursework 184
Concluding Statements 186

REFERENCES CITED....187

APPENDICES 194

APPENDIX A: Interview Protocol....195
APPENDIX B: Consent Form with IRB Approval 204

LIST OF TABLES

Table Page

1. Rubric for Analysis of Definition Items 80

2. Rubric for Analysis of Construction Items 81

3. Rubric for Analysis of Evaluation Items 85

4. Coding of David's Responses to Definition Items 115

5. Coding of Bob's Responses to Definition Items.....120

6. Coding of Bridget's Responses to Definition Items 122

7. Coding of Sam's Responses to Definition Items 125

8. Coding of Linda's Responses to Definition Items.....127

9. Coding for Construction Items.....143

10. Summary of David's Argument Evaluations.....150

11. Summary of Bob's Argument Evaluations 153

12. Summary of Bridget's Argument Evaluations.....157

13. Summary of Sam's Argument Evaluations.....158

14. Summary of Linda's Argument Evaluations 165

LIST OF FIGURES

Figure Page

1. Argumentation Framework 62

2. Construction Item 1 from Knuth (2002b) 65

3. Construction Item 2 Adapted from Stylianides & Stylianides (2009) 66

4. Construction Item 3 from Martin & Harel (1989) 67

5. Researcher Generated Responses for Evaluation Task 1 68

6. Researcher-generated responses to Evaluation Task 2 (Adapted from Martin & Harel, 1989; Knuth, 2002b) 70

7. David's Attempted Solution to the Angle Sum Problem 103

8. Bob's Written Part of the Solution to the Angle Sum Problem 105

9. Sam's Written Part of the Solution to the Angle Sum Problem 108

10. Linda's Written Part of the Solution to the Angle Sum Problem 110

11. Bridget's Written Part of the Solution to the Second Construction Item 136

12. Bridget's Initial Solution Attempt for the Calculus Item 136

13. Sam's Written Part of the Solution to the First Construction Item. 137

14. Sam's Written Part of the Solution to the Second Construction Item 138

15. Sam's Written Part of the Solution to the Calculus Item. 140

16. Linda's Written Part of the Solution to the Second Construction Item 141

17. Researcher Generated Responses for Evaluation Item 1. 145

18. Researcher Generated Responses for Evaluation Item 2 146

19. Diagram of Participants' Definition of Viable Arguments 176

ABSTRACT

This research investigated pre-service secondary teachers understanding of the term viable argument a new term introduced with the Common Core Standards. The research investigated how they define and understand viable argument, how they construct arguments, and how they evaluate mathematical arguments. The research was conducted using semi-structured interviews with five pre-service teachers that have completed the majority of their mathematical coursework. The results showed that the pre-service teachers compared viable arguments to mathematical proof and they recognized both similarities and differences between these terms. The participants were found to have the understanding of proof, reasoning, and argumentation that will be needed to implement the recommendations of the Common Core Standards.

CHAPTER 1

STATEMENT OF PROBLEM

Introduction

It is widely agreed that proof, reasoning, and mathematical argumentation are among the cornerstones of mathematical activity, both at the professional and the novice levels. Proof in mathematics has a number of roles and purposes, among which are validating results, promoting understanding, and explaining results (deVilliers, 1999). In an educational setting, proof, reasoning, and mathematical argumentation's main function is to explain and promote understanding of a mathematical object (Hanna, 2000a).

Nevertheless, these practices have not traditionally had a strong position in the mathematics classrooms in the U.S. (Hanna, 2000b), a claim supported by the results of the TIMMS videotape studies: In a random sample of 50 videotaped lessons of 8^{th} grade students (30 from 1995 and 20 from 1999), Jacobs et al (2006) found no lessons that involved proof. Furthermore, in the 1999 sample, none of the 20 videotaped lessons showed evidence of developing a rationale, making generalizations, or using counterexamples, by the teacher or by the students (p.20). In the United States, the practice of proof is typically introduced in high school geometry classes where students rigorously prove theorems they already know to be true, often in a strict two column format (Stylianides & Stylianides, 2009). The result is that proof and reasoning can be seen as a "teacher-game" (Reid & Knipping, 2010), an activity that earns grades and acceptance from the teacher but is otherwise useless.

Unfortunately, such misconceptions about proof are not only restricted to students. In one study, a participating teacher remarked that there is so much material to cover that proof is the first material to be dropped (Knuth, 2002b), thereby treating proof as a "topic of study rather than a means of coming to understand mathematics" (p. 420). The end result of such misconceptions is that students at all levels have a novice understanding of proof, reasoning, and mathematical argumentation (Reid & Knipping, 2010).

In recent years, standards policy writers and mathematics education researchers have emphasized mathematical proof, reasoning, and argumentation as important aspects of mathematics education that should be included at all grade levels (NCTM, 2000; Hanna, 2000; Conference Board of the Mathematical Sciences, 2001, 2012; Ball, Hoyles, Jahnke & Movshovitz-Hadar, 2002; Stylianou, Blanton & Knuth, 2009; Common Core State Standards, 2011). The National Council of Teachers of Mathematics (NCTM) states in their 2000 *Principles and Standards for School Mathematics* (the NCTM Standards) that "Reasoning and proof should be a consistent part of students' mathematical experience in prekindergarten through grade 12" (p. 56). The Conference Board of the Mathematical Sciences (CBMS) stated in 2001 that "mathematicians need to help prospective teachers develop an understanding of the role of proof in mathematics….Prospective teachers at all levels need experience justifying conjectures with informal, but valid arguments if they are to make mathematical reasoning and proof a part of their teaching. Future high school teachers must develop a sound understanding of what it means to write a formal proof" (p. 14). The Common Core State Standards

(CCSM) (2011) lists the ability to "construct viable arguments and critique the reasoning of others" as one of eight Standards for Mathematical Practice (p.6). With the exception of CBMS, these documents primarily address expectations for students in K-12 mathematics classrooms, but they have immediate implications for teachers as they are expected to facilitate the development of students' argumentation skills (CBMS, 2012). Teachers at both the elementary and secondary levels must be able to "make judgments about the logic used in arguments, even when the arguments are presented informally" (Yopp, 2010, p.411) and be able to provide students with possible avenues for improving an argument.

In the research on proof and reasoning, much attention has been given to the difference between deductive arguments and inductive arguments (Martin & Harel 1989; Knuth, 2002a, 2002b; Morris, 2007; Harel & Sowder, 2007; Stylianides & Stylianides, 2009). There are several important distinctions between inductive and deductive arguments. First, inductive arguments rely on one or more examples, whereas deductive arguments rely on a chain of logical assertions based on axioms, definitions, and previous results. Second, for mathematical statements over infinite sets, i.e. when there is an infinite number of possible cases that needs to be tested, inductive arguments are mathematically invalid in that they do not sufficiently demonstrate the truth of a claim because of the impossibility of testing all cases. On the other hand, deductive arguments can cover all cases, are mathematically valid, and are considered mathematical proof as long as the deductions in the argument are correct and logically valid. The main distinction between the two types of argument is therefore one of mathematical validity

based on the ability to account for all possible cases. Consequently, researchers have predominantly studied these arguments in terms of validity and called for measures that help students at all levels see the limitation of inductive arguments and the need for deductive arguments in mathematics (Harel & Sowder, 2007; Stylianides & Stylianides, 2009).

Although inductive arguments are generally not sufficient to prove a mathematical statement over infinite sets, some researchers (Balacheff, 1988; Rowland, 2002; Raman, 2003; Morris, 2007; Yopp, 2009) have identified a particular type of example-based argument, often referred to as generic example or generic proof, that has the ability to explain why a mathematical statement over an infinite set is true. These arguments contain key ideas (Raman, 2003) that provide a justification for the truth of the statements and that can potentially allow the given example to be generalized and formalized into a mathematical proof. It can be argued that this type of argument in their own right constitutes a form of mathematical reasoning (Staples et al., 2012), and it has been shown that professional mathematicians often rely on these types of arguments in their proof production (Raman, 2003). It is therefore important for teachers to be aware of both the mathematical and pedagogical merits of this type of argument as a precursor to mathematical proof (Rowland 2002; Yopp, 2010).

This research project investigates how pre-service secondary teachers, who have completed at least one course on mathematical proof, "construct viable arguments and critique the reasoning of others" as recommended by the CCSS. A first goal of the research is to explore how the pre-service secondary teachers understand and define the

notion of a viable argument and to determine what types of arguments they classify as viable. A second goal of the research explores the methods pre-service secondary teachers employ when they construct arguments in support of mathematical statements over infinite sets. A third goal is to explore how pre-service secondary teachers evaluate student arguments, what rationale lies behind their classification of student arguments as viable or non-viable, and whether they are able "to compare the effectiveness of two arguments, distinguish correct logic and reasoning from that which is flawed, and – if there is a flaw in the argument, explain what it is" (CCSS, 2011). In these explorations, of particular interest are pre-service teachers' use of key idea arguments and generic examples; if they distinguish between the merits of an empirical argument that can be generalized and an empirical argument that does not generalize; and whether they consider such arguments viable or not.

Background

Prior to the publication of the NCTM standards, research indicated that students had limited knowledge about mathematical reasoning, argumentation, and proof (Fischbein & Kedem, 1982; Martin & Harel, 1989; Harel & Sowder, 1998). At the same time, multinational mathematics assessments like the 1998 TIMMS study showed that American students underperformed in mathematics in comparison with other countries (Stigler & Hiebert, 1999). One of the key findings in the TIMMS video study conducted by Stigler and Hiebert was that American mathematics instruction to a large extent focused on memorization and procedure, whereas students in Japan and Germany were

more used to mathematics problems that induce critical thinking and require mathematical reasoning (Stigler & Hiebert, 1999; Jacobs et al., 2006).

In the period coinciding with and immediately following the publication of the NCTM Standards, the K-12 mathematics education community experienced a renewed interest in proof and reasoning. Articles were published about the role of proofs in mathematics (deVilliers, 1999) and in mathematics education (Hanna, 2000a); conceptions of proof and proof schemes (Harel & Sowder, 2007); the role of logic in proofs (Epp, 2003); the structure of proof production (Raman, 2003; Weber & Alcock 2004); teachers' and students' understanding of proof (Knuth, 2002a; 2002b; Selden & Selden, 2003; Riley, 2003); knowledge of proof for teaching and how to teach proof (Ball & Stylianides, 2008); proof-and-reasoning in school mathematics textbooks (Stylianides, 2009), and design experiments relating to proof (Stylianides & Stylianides, 2009). These studies have helped shape the understanding of educational difficulties associated with proof by identifying aspects of proof and reasoning that are problematic to learners. They have also highlighted important differences in both the usage and understanding of the terms "proof" and "reasoning" (Balacheff, 2002; Reid & Knipping, 2010) and the implications that this has for pedagogy (Ball & Stylianides, 2008; Stylianides, 2008).

More recently, standards documents and research have started focusing on informal reasoning and mathematical argumentation (CCSS, 2011), the structure of informal mathematical arguments (Walter & Barros, 2011; Yopp, 2011) and the value of informal reasoning and argumentation in mathematical proof (Rowland, 2002; Raman, 2003). The Common Core State Standards (2011), which has currently been adopted by

45 US states, includes eight Standards for Mathematical Practices, one of which is that students should "Construct viable arguments and critique the reasoning of others." Likewise, the President's Council of Advisors on Science and Technology (2011; 2012) calls for an emphasis on active learning techniques to build reasoning skills and for increased assessment of the type of reasoning required for college coursework. The main implication of the shift towards more informal ways of argumentation is that some of the rigor associated with the term "proof" is replaced by a greater focus on the process of discovering, proving and explaining mathematical results. In conversation, Dr. William McCallum, one of the lead authors of the Common Core State Standards, confirmed that the phrase "viable argument" entails a desired shift away from the rigorous connotations of the word "proof."

The essential message from proof research and national standards is that reasoning is considered a fundamental aspect of mathematics, and the processes associated with reasoning about mathematics are as important as mathematical content. Although researchers may interpret the terms "proof," "reasoning," and "argumentation" differently, they embrace a common goal: for students of mathematics at all levels to become proficient at making and justifying mathematical claims. The document that will guide curriculum and assessment for the foreseeable future, the Common Core State Standards, makes this goal clear: "Mathematically proficient students understand and use stated assumptions, definitions, and previously established results in constructing arguments. They make conjectures and build a logical progression of statements to

explore the truth of their statements…They justify their conclusions, communicate them to others, and respond to arguments of others" (2011, pp. 6-7).

Problem Statement

As mentioned earlier, research has shown that students at all levels have inadequate understanding of mathematical proof, reasoning, and argumentation (Balacheff, 1988; Martin & Harel, 1989; Healy & Hoyles, 2000; Knuth, 2002a, 2002b; Riley, 2003; Selden & Selden, 2003). In particular, several studies have exposed students' insufficient understanding of inductive and deductive argumentation and questions relating to mathematical validity (Morris 2002, 2007). Research has shown that students readily accept examples as proofs and also "prove" statements using empirical evidence in situations where examples cannot exhaust the number of possible cases. Researchers highlight the importance of students recognizing the limitations of empirical arguments and the necessity of deductive arguments (Martin & Harel 1989; Healy & Hoyles, 2000; Knuth, 2002; Harel & Sowder, 2007; Morris, 2002, 2007; Reid & Knipping, 2010) and there is a call among proof researchers to develop methods that help facilitate this recognition: "Ultimately, the goal is to gradually develop an understanding of proof that is consistent with that shared and practiced in contemporary mathematics" (Harel & Sowder, 2007).

The task of developing these understandings in students will eventually become the responsibility of classroom teachers. Hence, it is essential to know if teachers are

adequately prepared to help facilitate students' transition from empirical to deductive modes of argumentation.

Purpose of the Study

This study investigates pre-service secondary teachers' understanding of the term "viable argument" as well as the strategies they employ to "critique and construct viable arguments," one of the eight Standards for Mathematical Practice embedded throughout the Common Core State Standards. If students are to be proficient at this, it is imperative that their teachers are as well. The study will investigate the proof, reasoning, and argumentation skills of pre-service secondary teachers enrolled in a teacher preparation program that closely follows the recommendations set forth by the CBMS for the Mathematical Education of Teachers (2001, 2012) and that incorporate the standards and practices from the Common Core State Standards in classroom instruction.

A series of task-based interviews focuses on three key components supporting the practice of "constructing viable arguments and critiquing the reasoning of others." First, how do pre-service teachers define or understand what is meant by a viable argument, and how do they distinguish between proof, viable arguments and non-viable arguments? Second, are they able to differentiate between empirical evidence used in a constructive fashion and empirical evidence that cannot be extended to a more formal argument? Third, what strategies do pre-service teachers employ when constructing arguments, empirical or deductive?

Research Questions

The current study will focus on the interpretation, construction, and evaluation of viable arguments, all with respect to pre-service secondary teachers' understanding of how to "construct viable arguments and critique the reasoning of others" after having completed the majority of a teacher preparation program closely aligned with the CBMS's recommendation for the Mathematical Education of Teachers. A series of task-based interviews was conducted to explore the following research questions.

1. How do pre-service teachers define the notion of "viable argument"? Is their notion consistent with "viable argument" as defined by the Common Core Standards?
2. What strategies do pre-service teachers employ when asked to construct viable arguments?

   2a: Do they build a logical progression of statements to explore the truth of conjectures?

   2b: Do they use examples constructively in their arguments?
3. What characteristics do pre-service teachers assess to critique the reasoning of others?

   3a: Do they distinguish between correct logic and logic that is flawed?

   3b: Do they look for ways to clarify or improve the arguments of others?

   3c: Do they recognize and value constructive uses of examples?

Significance of the Study

Current research demonstrates the need for improved understanding of proof, reasoning, and argumentation among U.S. students, and current standards reflect this need by calling for more and improved instruction in this area. The development of these skills begins in the elementary classroom and continues throughout the grades. The responsibility for developing students' argumentation and reasoning skills does not lie solely within the K-12 system, but also with teacher preparation programs and higher education. It is imperative that pre-service teachers understand the importance of proof, reasoning, and argumentation and also be adequately prepared to teach these topics. Consequently, courses for future teachers must incorporate these skills, not only in terms of developing teachers' own knowledge, but also in developing their proficiency in assessing their students' proof, reasoning, and argumentation skills and suggesting improvements to student arguments. This means that teacher preparation courses and materials must incorporate mathematical argumentation, reasoning, and proof throughout the curriculum. This is highlighted by CBMS's Mathematical Education of Teachers (2001) which states that high-school teachers must "develop a sound understanding of what it means to write a formal proof" (p. 14) and be able to use "algebraic reasoning effectively for problem solving and proof" (p. 40) in order to teach high school curricula effectively.

Several studies have shown clear deficiencies in the understanding of proof and reasoning among teachers and students alike. In particular, pre-service teachers have been shown to possess only rudimentary understanding of proof, reasoning and argumentation

and their importance in mathematical endeavors. Research has shown that pre-service elementary teachers accept empirical evidence as proof and use empirical evidence to prove inexhaustible mathematical statements (Stylianides & Stylianides, 2009; Reid & Knipping, 2010). Similar results have been documented among pre-service secondary teachers (Reilly, 2003) and in-service secondary teachers (Knuth, 2002a, 2002b). These studies all show that the empirical-deductive distinction is a topic that continues to cause problems for pre-service teachers at all levels.

However, the distinction between empirical and deductive reasoning has been exclusively studied in terms of proof, which means that a sharp line is drawn between the validity of the two types of argumentation and research has tended to treat inductive evidence as insufficient. Furthermore, pre-service teachers' and students' reliance on empirical evidence when constructing proof is often interpreted by the researchers as a lack of understanding—that is, as a belief that it is possible to prove a statement that covers infinitely many cases by using examples. However, another interpretation is also possible; it is entirely possible that the reliance on empirical evidence is evidence of an insufficient mathematical skill set (Stylianides & Stylianides, 2009). Students may simply have no other strategy for how to prove a statement and therefore resort to empirical evidence although they are aware of the limitations associated with this.

Some researchers have pointed out how empirical evidence can provide a basis for deductive argumentation and formal proof (Rowland, 2002; Yopp, 2009), and how professional mathematicians often rely on informal argumentation and empirical evidence prior to producing a valid proof (Raman, 2003; Weber & Alcock, 2004).

However, pre-service teachers' proficiency in emulating these practices and if they recognize how empirical evidence can be used as a precursor to more general arguments have not been studied. Similarly, if pre-service teachers' distinguish between empirical evidence that can be generalized and empirical evidence that is not generalizable has not been studied. There is therefore a need for research that aims to measure whether pre-service teachers are able to identify generalizing features of empirical arguments. If teachers are to help guide students towards more secure forms of validation, they need to be able to identify positive aspects or key ideas in students' arguments and explain how these aspects and ideas can be used to further improve those arguments. This is also essential if teachers are to help students critique the reasoning of others.

Limitations of the Study

Certain factors limit the scope and the ability to generalize the results of this study. First, the sample size had to be kept small to allow for an in-depth focus on student thinking and argumentation. This reduces the potential for extending the findings to other populations. Second, the subjects chosen for study are enrolled in a well-defined teacher preparation curriculum, which further limits how the results can be applied. Third, there is potential for bias as the researcher has extensive knowledge of the literature on proof and must be careful not to project that knowledge on the study participants. Fourth, the researcher does not have experience teaching the proof-based courses designed for pre-service secondary teachers in this program, and must guard against forming inaccurate

expectations. Strategies for increasing generalizability and reducing bias are discussed in greater detail in Chapter 3.

CHAPTER 2

REVIEW OF LITERATURE

Introduction

The Common Core State Standards (2011) lists students' ability to "construct viable arguments and critique the reasoning of others" as one of eight Standards for Mathematical Practice. This implies that teachers must be able to determine what constitutes a viable argument and have a clear image of which features of an argument are important. It also implies that teachers must be able to critique the reasoning of students in the context of what constitutes a viable argument (CBMS, 2012). The focus of this study is to investigate how pre-service secondary teachers interpret the term "viable argument" and investigate how they construct viable arguments and critique the reasoning in arguments.

This chapter starts with a review of how proof, reasoning, and argumentation have been used in the research literature and some difficulties arising from this use. Central to this review are the roles of proof, reasoning, and argumentation in mathematics and the teaching of mathematics, as well as how researchers envision the introduction of proof, reasoning, and argumentation into the K-12 school curriculum. The second part of the chapter discusses research on conception and understanding of proof, reasoning, and argumentation; differences between inductive and deductive reasoning and misconceptions about these differences; and the role of examples in the process of

proving and in the teaching of mathematics. The chapter will conclude with a review of articles relevant to the methodology for this study.

The Language of Proof, Reasoning and Argumentation

Definition of Terms

In the following, when referring to research conducted by others I will adopt the same terminology used by those conducting the research. As a consequence, there will be some blurring of distinctions between the terms proof, reasoning, and argumentation as their usage in the literature has considerable overlap. In many cases, what is described in terms of proof in the literature immediately extends to reasoning and argumentation. For instance, when verification is described as a role and purpose of proof, it is understood that this can by extension be regarded as a role and purpose of mathematical argumentation. In fact, researchers have shown how professional mathematicians often do not require formal proof as verification of the truth of a statement, but are content with an argument that describes the logical underpinnings of a proof or an informal idea that can be extended to a proof (Raman 2003).

In reference to my own research, I will use the term *mathematical argument* to describe a justification of a mathematical statement. In the terminology of Toulmin's argument analysis (Reid & Knipping, 2010), to qualify as a mathematical argument three components must be present: a claim, data, and a warrant. The claim is the mathematical result or conjecture in question, the data consist of the information that supports the claim, and the warrants provide the link between the data and the claim, i.e., they explain

why and how the data supports the mathematical claim. Under this definition, a mathematical argument can be based on deductive logic or not, it can be correct or incorrect, and it can be valid or invalid.

I interpret *reasoning* as the process of producing a mathematical argument. Reasoning is thus the cognitive process where the person making the argument chooses to pursue certain avenues of justification whereas others are discarded. I will use *proof* to mean a logically valid and correct mathematical argument consistent with how the term is used in undergraduate and graduate mathematics classes. A combination of these terms—proof, reasoning, and argumentation—will be used when I am referring to the literature in general. The reason for this is twofold. First, I see the three terms as interconnected. Second, I believe that all three terms describe a common goal; to have students make mathematical claims and be able to mathematically back up these claims.

Of particular interest to this research is the term *viable argument* used by the Common Core Standards for Mathematical Practice. The CCSS does not provide an exact definition of viable argument; instead its meaning is embedded in the following quote:

> **3. Construct viable arguments and critique the reasoning of others.** Mathematically proficient students understand and use stated assumptions, definitions, and previously established results in constructing arguments. They make conjectures and build a logical progression of statements to explore the truth of their conjectures. They are able to analyze situations by breaking them into cases, and can recognize and use counterexamples. They justify their conclusions, communicate them to others, and respond to the arguments of others. They reason inductively about data, making plausible arguments that take into account the context from which the data arose. Mathematically proficient students are also able to compare the effectiveness of two plausible arguments, distinguish correct logic or reasoning from that which is flawed, and—if there is a flaw in an argument—explain what it is. Elementary students can construct arguments using concrete referents such as objects, drawings, diagrams,

> and actions. Such arguments can make sense and be correct, even though they are not generalized or made formal until later grades. Later, students learn to determine domains to which an argument applies. Students at all grades can listen or read the arguments of others, decide whether they make sense, and ask useful questions to clarify or improve the arguments. (CCSS, 2011, pp.6-7)

I will use the term *viable argument* in a manner that is consistent with the above description, but with an additional requirement that mathematical statements over infinite sets must contain a feature that is generalizable; that is, a feature that if extended explains why the result holds true for all cases.

Vocabulary

One of the main difficulties in reviewing the research on proof, reasoning, and argumentation relates to vocabulary. As pointed out by Reid and Knipping (2010) the vocabulary used is hopelessly inconsistent; sometimes the same term is used differently from one research article to the next, whereas in other research reports different terms are used to describe the same phenomenon. One example of the type of inconsistencies Reid and Knipping allude to is found in the terms *generic example* (Balacheff, 1988), *generic proof* (Rowland,2002), *single case key idea* (Morris, 2007), *key idea* (Raman, 2003), and *pre-formal proof* (Blum & Kirsch, 1991). These terms can all be used to describe the same type of mathematical argument, an argument that consists of a single example that contains the logic sufficient to construct a mathematical proof.

The terms “proof,” “reasoning,” and “argumentation” are also used differently by different authors. For instance, some researchers claim that there is an ontological difference between “mathematical proof” and “argumentation” since the term argument

implies a kind of reasoning that is opposed to proving (Reid & Knipping, 2010, pp. 153-158). Among these researchers, Duval claims that informal argumentation actually impedes the learning of mathematical proof (in Durand-Gurrier et al., 2012). In Duval's opinion, in an argument conclusions based on inferences and shared knowledge are continuously reinterpreted based on their semantic content. On the other hand, in a mathematical proof, the conclusions reached are reached from mathematical deductions and the focus is on the operational status of the content rather than the semantic interpretation (Durand-Gurrier et al., 2012).

Other researchers have a more pragmatic view of mathematical argumentation and see it in terms of problem solving. This view is not so much concerned with ontological differences between proof and argumentation, instead they claim that "argumentation can be used as an effective basis for mathematical classroom discussions concerning not only mathematical reasoning in general but also the rules for mathematical proof and the mathematical practices that play important roles in proof production" (Durand-Gurrier, 2012, p.354). With this approach, mathematical argumentation can be regarded as a continuum of more or less sophisticated mathematical arguments. From a pedagogical standpoint, if one thinks of mathematical argumentation as a continuum, a formal mathematical proof can be seen as the most sophisticated type of mathematical argument and the goal towards which teachers should lead their students.

Another view encountered in the literature is that the term "proof" can be defined in a manner that makes it suitable for K-12 mathematics (Stylianides, 2007a; Schifter, 2009). This view stresses the social aspect of a proof by claiming that what is acceptable

as a proof is dependent on the classroom community. In this view, the definition of proof becomes dynamic and context-dependent, and an argument that is accepted as proof in a lower-grade classroom is not necessarily accepted in a higher-grade classroom (and vice versa) as the mathematical proficiencies of students differ at different grade-levels.

There are even differences in how the different terms are used within the same research article. Reid and Knipping (2010) describe how the terms *proof* and *prove* are used in at least three very distinct ways within the same article: to refer to a concept, an object, or a process (p. 28). Add to this that there is a multitude of different terms encountered in the literature; for instance, some make distinctions between the terms *proof* and *mathematical proof*, the former being an argument accepted by a community whereas the latter is an argument accepted by mathematicians (Martinez, Brizuela, & Superfine, 2011). Despite the confusion that may be caused by the plethora of terms and perspectives on these terms, the mathematics education community seems to converge on the underlying importance of proof, reasoning, and argumentation. There is a common call among researchers and standards documents—for students at all grade levels to form mathematical conjectures, construct mathematical arguments, and be able to justify their conclusions. This common goal takes precedence over the terminology used in order to achieve this goal.

An Overview of Proof

Proof Categories

Several researchers have classified different categories of proof based on the type of validation methods encountered in the proving process (Balacheff, 1988; Martin & Harel, 1989; Blum & Kirsch; 1991, Harel & Sowder 1998, 2009; Waring, 2000; Stylianides & Stylianides, 2009). These categories typically rely upon an experimental approach where the processes that govern conviction about the validity of a particular mathematical statement are studied (Balacheff, 1988). The categories described are always hierarchical where certain types of arguments are considered better than others. In the following pages, two different and well-known categorizations of proof will be discussed: Balacheff's levels and types of proofs (1988) and Harel & Sowder's proof schemes (1998, 2009). These two categorizations are arguably the most influential in the research on proof, reasoning and argumentation, coining terminology such as *naïve empiricism* (Balacheff, 1988) and *authoritative proof scheme* (Harel & Sowder, 2007) that have been assimilated into the canon of research on proof, reasoning, and argumentation. It is important to note that these levels all use the term *proof* although several of the argument categories described cannot be classified as proof. Balacheff (1988) explicitly states that the use of the word proof stems from the belief of the student that the argument he or she has produced is a proof.

Both categorizations provide useful insight into the type of student arguments teachers can expect to encounter in the classroom. Additionally, since they are hierarchical, the categories provide useful insight into what constitutes "the mathematical

horizon" (Hill, Ball & Schilling, 2008) in terms of proof, reasoning, and argumentation; that is, the levels of knowledge and skill in these areas that teachers are expected to help their students attain.

Levels of Proof and Types of Arguments

Balacheff (1988) distinguished between two types of proofs, *pragmatic* and *conceptual*. The distinction arises from whether or not the argument relies on concrete representations such as physical objects, diagrams, or pictures. A pragmatic proof relies on such concrete representations, whereas a conceptual proof does not. Instead a conceptual proof uses formulations of mathematical properties and the relationship between these, often abstractly. Among various types of pragmatic and conceptual proof, Balacheff (1988) focuses on four types of arguments that are particularly prevalent in the cognitive development of proof: *naïve empiricism*, the *crucial experiment*, the *generic example*, and the *thought experiment*.

By *naïve empiricism*, Balacheff is referring to arguments that assert the truth of a statement based on checking a small number of particular cases. For example, a student that verifies the truth of a statement by checking only the first five cases exhibits the characteristics of naïve empiricism. Balacheff further exemplifies naïve empiricism with the following classroom example. By checking three different polygons, a square, a hexagon, and an octagon, two students concluded that the number of diagonals in a polygon is found by dividing the number of vertices by two. When one of the students later came up with a counterexample by checking a pentagon, they simply devised another formula for polygons with an odd number of sides. This amendment was

accepted by additionally checking a heptagon. Consequently a total of five examples were sufficient to find a general assertion about the number of diagonals in a polygon.

A slightly more sophisticated argument is the *crucial experiment*, where the examples are picked strategically to verify the truth of a statement. An example of the crucial experiment would be a student testing an assertion who, in addition to checking the first few cases, also intentionally checks special cases such as a large number or a prime number. When the assertion holds true for these special cases, the general assertion is accepted as true. The important difference between these two stages is that students at the crucial experiment stage exhibit signs of attempting to generalize by intentionally testing special cases. It is also important to note that the intent of the student is a defining characteristic for the crucial experiment; although the method is still mathematically invalid, there is at least recognition on the part of the student that patterns may not hold for all numbers and the conjecture is therefore strategically tested.

The next stage in Balacheff's cognitive development of proof is referred to as the *general example*. An argument belonging to this class can be seen as a blueprint of a valid proof, in that it uses a specific example in a deductive manner. The example thus serves as a representative of its class, and the argument includes the structure and logical implications of a valid proof. To exemplify, Balacheff uses the following example taken from Bezout's *Notes on Arithmetic* (1832):

> The remainder on dividing a number by 2x2 or 5x5 is the same as the remainder on dividing the number formed by the rightmost two digits by 2x2 or 5x5….
> To fix these ideas, consider the number 43728 and the divisor 5x5. The number 43728 is equal to 43700+28. However, 43700 is divisible by 5x5, because 43700 is the product of 437 and 100, and as 100 is 10x10 or 5x2x5x2, the factor 100 is

> divisible by 5x5. The remainder on dividing 43728 by 5x5 or 25 is therefore the same as that on dividing 28 by 25. (in Balacheff, p.219)

The main feature of this argument classifying it as a generic example is the de-contextualization of the example used. There is nothing special about the numerical example given in the argument, 43728, since the exact same argument can be applied to any multi-digit number. The truth of the statement does not rest on the example used, but rather on logical deductions used in the argument. The example becomes a vehicle for showing why the statement is true.

The last of Balacheff's categories is called the *thought experiment*. At this level, students draw logical conclusions based on the mathematical properties and relationships present in the situation. Since this level signifies that the students have been able to distance themselves from concrete representation by relying on abstractions, this is the only of the four levels that take place in the conceptual rather than the pragmatic realm. As an example of a thought experiment, Balacheff provides the following argument from two students working on finding the relationship between the number of diagonals and the number of vertices in a polygon: "In a polygon if you have x vertices the number of diagonals which go from one point will be x-3 because there is one point you are leaving from and the two points which join it to the polygon" (1988, p. 225).

The important feature of this argument that classifies it as a thought experiment is that the two students are not referring to one or several special cases in their argument. Initially these students worked with a hexagon in a manner consistent with a generic example, but they were able to abstract their result to embrace all polygons. It is also important to note that it is the method of argumentation and not the validity of the

argument or correctness of the final answer that is important to Balacheff, since the argument above fails to take into account that there are x vertices, which means that the correct formula for the number of vertices in a polygon is $s(x) = \frac{x(x-3)}{2}$.

These four categories constitute a progression of levels of increasing sophistication, and Balacheff explicitly states that there is a fundamental difference between the first two, which rely on examples only, and the latter two, which rely on logical deduction:

> For the generic example and the thought experiment, it is no longer a matter of 'showing' the result is true because 'it works'; rather, it concerns establishing the necessary nature if its truth by giving reasons. It involves a radical shift in the pupils' reasoning underlying these proofs (p.218).

Although Balacheff's work has been influential in the study of proof, reasoning, and argumentation, using Balacheff's taxonomy of proof in order to categorize student work has proved difficult. The different categories can only be applied based on the intentions and cognitive processes in the students, processes which very often are difficult to ascertain from written work only. Varghese (2011) claims that in particular, the distinction between naïve empiricism and the crucial experiment can be difficult to distinguish, primarily because it is hard to determine whether or not a special example has been picked strategically without explicitly asking the person supplying the argument. Based on examination of research articles, Varghese also questions how prevalent instances of the generic example will be in student work (p. 185). Despite these reservations, Balacheff's taxonomy of proof provides a useful framework for describing the types of arguments likely encountered in student work.

Proof Schemes

Harel and Sowder (1998, 2009) define proving as the process of removing doubt about the truth of an assertion, where an assertion is a conjecture (i.e., the truth-value is not known to be individual making the assertion) or a fact. A conjecture becomes a fact when the individual becomes certain about its truth. Proving has two sub-processes: ascertaining, which is removing one's own doubt; and persuading, removing the doubt of others. Based on this they define a proof scheme in the following manner: "A person's (or a community's) proof scheme consists of what constitutes ascertaining and persuading for that person (or community)" (2009, p.809). A proof scheme is therefore the methods individuals use in order to convince themselves and others about mathematical statements. Even though Harel and Sowder explicitly use the terminology of proof and proof schemes, a proof scheme essentially describes the types of argument that a person finds convincing.

Based on their definition of a proof scheme, Harel and Sowder outline a hierarchy of argumentation strategies that represent cognitive stages in students' mathematical development. Their categories consist of three main categories or classes. First is the *external conviction proof schemes* class, where students are convinced by external factors. Such external factors may be an authority such as a teacher or a textbook, or the convincing appearance of an argument such as the two-column proof format. Second, the *empirical proof schemes* class consists of two sub-classes, the *inductive proof scheme* class and the *perceptual proof scheme* class. These two classes are similar to the first three categories offered by Balacheff (1988): the inductive proof scheme class is similar

to naïve empiricism, and the perceptual proof scheme class is similar to the crucial experiment or the generic example in that conviction is based on the use of one or more strategic examples. The last category is the *deductive proof scheme* class. In this category, Harel and Sowder recognize two sub-classes; *transformational* and *axiomatic* proof schemes, where students are convinced by arguments that include generality, operational thought, and logical inference. The deductive proof scheme class is therefore similar to Balacheff's fourth category, the thought experiment.

As in Balacheff's taxonomy a student's cognitive stage is measured by the type of argument he or she produces and the categories are viewed as hierarchical, where the deductive proof scheme class is the most sophisticated cognitive level. Harel and Sowder stress that despite the intentionally subjective and student-centered definition of a proof scheme, "the goal of instruction must be unambiguous–namely, to gradually refine current students' proof schemes toward the proof scheme shared and practiced by contemporary mathematicians" (p. 809). They also state that the main question for mathematics educators is how teachers can utilize their students' existing proof schemes, which research has shown to be largely external and empirical, to help them reach the deductive proof scheme level (p.817). If they are to do so, teachers must be able to recognize generalizing features in empirical arguments and find avenues using the features of an empirical argument to make a deductive argument.

The Empirical Deductive Distinction

The distinction between the empirical and the deductive is important in the study of proof, reasoning, and argumentation. This is evident in the proof taxonomies of

Balacheff (1988) and Harel and Sowder (1998, 2009) and from the numerous research studies on the distinction between empirical and deductive argumentation (Martin & Harel, 1989; Healy & Hoyles, 2000; Morris, 2002, 2007; Stylianides & Stylianides, 2009). The main difference between empirical and deductive justification relies on the method employed to validate an argument. With empirical reasoning, one or more examples are employed to validate a conjecture. Empirical methods are widely used in the natural sciences as experimental tests of theoretical hypotheses, but for mathematicians empirical tests are insufficient to validate statements where infinitely many cases would need to be tested. Since empirical evidence can never exhaust the possible cases, empirical methods are logically invalid and are not sufficient to prove a mathematical statement. To give an example, it is impossible to prove through examples alone that an even number added to an even number equals an even number, since there are infinitely many cases that would need to be checked.

As opposed to empirical justifications, deductive justifications rely on logical inferences and are considered valid mathematical arguments. Therefore, to prove a conjecture on an infinite set, the field of mathematics typically requires some sort of deductive argument. For instance, to prove the conjecture that the sum of two even numbers is even, one can argue the following: Even numbers are on the form 2n where n is a natural number. Then the sum of two even numbers can be expressed as 2n+2m where n and m are natural numbers. Finally 2n+2m=2(n+m), which is always an even number since n+m is also a natural number. The above argument is general in nature

since it does not depend on which natural numbers are chosen for n and m. It therefore covers all cases, and there is no need to further check the result by specific examples.

From a mathematical standpoint, the difference between empirical and deductive justification is not just a difference in validation method, but also a difference in logical validity. Deductive arguments are logically valid arguments since they cover all possible cases and when a statement has been proven deductively its generality has been confirmed and there is no need to further check the statement. Empirical evidence can only establish the truth for specific cases, but cannot be applied to prove a statement when there are an infinite number of cases to check. Empirical justifications on statements pertaining to infinite sets are therefore considered logically invalid modes of argumentation.

From a pedagogical standpoint, the situation is different. Rowland (2008) argues that the role of empirical evidence in teaching is essentially two-fold. First, examples provide particular instances of something general. We teach procedures by providing examples—that is, specific instances of the general procedure. This teaching practice is commonplace and it aims to simplify general concepts and typify general procedures. Rowland argues that "in the case of concepts, the role of examples is to provoke or facilitate abstraction" (p. 150). This role of examples is in essence inductive, to give specific examples of general concepts. The second role of examples is essentially illustrative and practice-oriented, where students are intended to understand and retain a procedure through repetition of that procedure.

The Role of Examples

Until recently, the role of examples in proof, reasoning, and argumentation has not received much attention. Most research points to the fact that examples provide empirical evidence which is logically invalid and cannot be regarded as mathematical proof. Consequently, research has focused on whether a given community fully understands the difference between empirical and deductive justifications, and how to teach and help students understand the difference between the two (Morris, 2002; Harel & Sowder, 2009; Stylianides & Stylianides, 2009b).

Even though the push is away from empirical justification towards deductive proof schemes (Harel & Sowder, 2009), most researchers also value the use of examples in mathematical exploration. Stylianides & Stylianides (2009b) state that empirical exploration can "help students organize mathematical observations into meaningful mathematical generalizations" (p. 315). They further state that empirical explorations can help gain insight into how to prove the generalizations. Yet, empirical evidence should never be substituted for logically valid arguments since this is inconsistent with current mathematical methods. Morris (2002) expresses similar sentiments when she rates example-based inductive arguments and single-case key idea inductive arguments (her term for a general example) as mathematically invalid and therefore not proof, but, at the same time, Morris acknowledges the possibilities for single-case key-ideas in explaining why a statement is true, and speculates that generic examples can be used to teach generalization and valid modes of proof. Other researchers take a different stand: Because of its deductive nature, Martin and Harel (1989) classify the general example

used in their study as a deductive argument. Based on these different viewpoints, whether one should classify a general example as inductive because it fails to generalize, or as deductive because of its deductive nature and the fact that it is a valid proof camouflaged as an example can be debated.

Argumentation in the K-12 Classroom

Roles and Purposes of Proof

Building on an earlier classification by Bell, deVilliers (1999) outlines six major roles of proof in mathematics:

- *verification* (concerned with the truth of the statement)
- *explanation* (providing insight into why it is true)
- *systematization* (the organization of various results into a deductive system of axioms, major concepts and theorems)
- *discovery* (the discovery or invention of new results)
- *communication* (the transmission of mathematical knowledge)
- *intellectual challenge* (the self-realization/fulfillment derived from constructing a proof) (p. 2).

It is understood in this classification that these roles are not supposed to be considered mutually exclusive; also, and although deVilliers limits his discussion to proof, many of the same roles can be extended to mathematical reasoning and argumentation as well.

In addition to the roles and purposes of proof outlined by deVilliers, proof and argumentation take on additional roles when considered in an educational setting:

> Proof is central to mathematics and as such should be a key component of mathematics education. This emphasis can be justified not only because proof is at the heart of mathematical practice, but also because it is an essential tool for promoting mathematical understanding (Ball, Hoyles, Jahnke, Movshovitz-Hadar, 2002, p. 907).

This quote offers two important additions to deVilliers' list of proof roles. First, the process of proving represents authentic mathematics—it is what mathematicians do. Through proving mathematical statements, students gain insight into the field of mathematics. Second, proof and reasoning promotes mathematical understanding, because they provide the rationale behind mathematical results and procedures. These important roles are echoed by Yackel (2002): "Issues of what it might mean to explain and justify, and how students might learn to do so, are of primary concern to researchers and classroom teacher alike" (p.423). Similarly, Hanna (2000a) states that the most important part of proof in the classroom is to promote understanding of the mathematics presented and that the challenge is to find ways to effectively use proof for this purpose: "Proof can make its greatest contribution in the classroom only when the teacher is able to use proofs that convey understanding" (p. 7). In this sense, Hanna suggests that even though it is important for students to understand the nature of deductive reasoning, it is even more important to gain mathematical understanding from proofs. It is understood that in mathematics teaching, proof primarily becomes a vehicle for understanding rather than a verification tool.

Definitions of Proof for School Mathematics

The notion of mathematical proof as a social construct and hence context dependent, has widespread support within the philosophy of mathematics (Stylianides

2007b). To a large extent, the social constructivist view of proof stems from the fact that a mathematical proof builds on definitions, logical deductions, and prior results. Definitions and prior results are dependent on a student's knowledge base, recognizing that "students' construction of new knowledge is based on what they already know" (Harel & Sowder, 2007, p.807). The students' knowledge base again depends on both individual and social factors, and it therefore follows that both definitions and prior results are subjective and context dependent.

The idea of proof as a social construct has also influenced mathematics education research and contributed to alternate definitions of proof specifically geared towards mathematics education. While recognizing the traditional definition of proof, Harel & Sowder (2007) define a proof as "what establishes truth for a person or a community" (p. 806) and hence argue that "proof" should be interpreted subjectively. The benefit of such a definition is that it allows for proof to permeate the curriculum through all grade levels.

Stylianides and Ball (2008) define a proof as a mathematical argument that fulfills three criteria:

> (i) it uses statements accepted by the classroom community (*set of accepted statements*) that are true and available without further justification.
> (ii) it employs forms of reasoning (*modes of argumentation*) that are valid and known to, or within the conceptual reach of, the classroom community; and
> (iii) 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. (p. 309)

They argue that this definition consists of three main components: the set of acceptable statements, the modes of argumentation, and the modes of argument representation. They claim their definition seeks to strike a balance between two often competing aspects of proof in school mathematics: the view of mathematics as a

discipline and the view of students as mathematical learners. This balance is achieved by requiring that a proof uses true statements and valid modes of reasoning—characteristics important for a mathematical argument to be considered a mathematical proof—but at the same time it can be modified in terms of the classroom community, thereby making the definition of proof dependent on the mathematical community in which the argument is presented.

Similarly, Schifter (2009) contends that children in elementary school, albeit capable of justifying claims of generality, are still coming to terms with the four basic arithmetic operations and cannot be expected to construct proofs or fully understand the concept of a mathematical proof. Schifter claims that arguments from representation, whether these are physical objects, pictures, diagrams or a story context, can provide a powerful way for elementary students to convey ideas of generality. She offers the following three criteria for what can be regarded as a *representation-based proof* in the elementary grades:

1. The meaning of the operation(s) involved is represented in diagrams, manipulatives, or story contexts.
2. The representation can accommodate a class of instances (for example whole numbers).
3. The conclusion of the claim follows from the structure of the representation. (p.76)

The different definitions in this section highlight how mathematics education researchers have incorporated a social dimension and subjectivity into the realm of mathematical proof by defining what constitutes a proof as a function of its accessibility to the audience. The benefit of this is that it allows proof to be taught at all grade levels and to permeate all branches of mathematics, as opposed to being confined to high-school

geometry. It also removes some of the ritualistic or formalistic aspects of the teaching of proof associated with the two-column format (Weiss, Herbst, & Chen, 2009).

Problems with Definitions of Proof for School Mathematics

Even though the different social definitions of proofs have certain merits as outlined above, there are also certain drawbacks associated with these definitions. One of these is that the above definitions contradict one of the main tenets of mathematical definitions, that they be precise and unambiguous. By making the definition of *proof* depend on the audience the definitions become imprecise and it can be open to interpretation whether an argument is a proof or not.

Using a social dimension as a guideline for whether an argument qualifies as proof or not may also lead to the potentially problematic situation of an argument being rejected by the classroom community because it is too sophisticated for that community (Stylianides 2007b, 17). This problem can be solved when the teacher acts as the classroom representative of the mathematics community by intervening and helping students develop and socially share the argument put forth. However, in order to do so effectively, the teacher must have a firm grasp not only of content knowledge but of what Shulman (1986) coined *pedagogical content knowledge*, which means the teacher simultaneously needs to hold two different justification schemes in regards to an argument. First, the teacher needs to be aware of whether or not the argument is a mathematical proof. Second, the teacher needs to be aware if the argument is acceptable

to the community of learners and can be accepted as a proof by the students in the classroom.

Another problematic aspect inherent in any definition of proof that emphasizes the social dimension is that the definition often relies heavily on the proof's ability to convince (Raman, 2003; Harel & Sowder, 2007). Creating conviction about the truth of a statement is undeniably an important aspect of proof in mathematics, but conviction often comes prior to the proof itself, so conviction is not in itself sufficient for proof (Hanna, 1998). Furthermore, it is possible to be convinced and incorrect, and an initial conviction may be stronger than any valid arguments contrary to the initial conviction. For instance, Williams (1991) documented how a student was willing to deny her own physical experience of a train stopping at a station because of her concept image (Tall & Vinner, 1981) of limits convincing her that a function cannot reach its limit. In another study, an in-service elementary teacher did not give up her incorrect conviction that $.999\ldots \neq 1$ despite several arguments given to the contrary (Yopp, Burroughs & Lindaman, 2011). As these examples show, basing proof on the ability to convince is at best problematic.

Proof versus Viable Argument

It is noteworthy that the term *viable argument* as used in the Common Core State Standards emphasizes the same characteristics of an argument as do Stylianides and Ball (2008) and Schifter (2009) in their respective definitions of proof for K-12 mathematics. A viable argument uses mathematical reasoning, it contains generalizing features, it may be informal or representation-based (Schifter, 2009), and it is dependent on the classroom community. However, *viable argument* avoids all the connotations of formalism and

rigor that the term *proof* conjures in mathematics. Rather than focusing on whether or not an argument can be classified as a proof, the emphasis is on modes of argumentation true to the field of mathematics. At the same time, viable arguments take the student as a learner of mathematics into consideration, precisely the balance Stylianides and Ball try to achieve in their definition of a proof for school mathematics.

Teaching Proof, Reasoning, and Argumentation

Teachers' Understanding of Proof and Argumentation

Although most research about the conception of proof, reasoning, and argumentation has focused on students, a few key studies have been conducted on teachers and pre-service teachers' conception of proof and ability to prove. Martin and Harel (1989) showed that pre-service elementary teachers accepted both inductive (empirical) and deductive arguments as proof, noting that "80% of the students gave a high validity rating to at least one inductive argument and over 50% gave a very high validity rating to at least one inductive argument" (p. 47). Furthermore, acceptance of inductive and deductive arguments was not mutually exclusive, which led the researchers to hypothesize that pre-service elementary teachers simultaneously exhibit two different proof frames, one inductive and one deductive. The inductive proof frame is a result of their experience from everyday life, whereas the deductive proof frame is a result of mathematics instruction. Martin and Harel also hypothesized that instruction in deductive reasoning does not automatically replace the propensity to use inductive reasoning. However, changing the terminology from *proof* to *viable argument* may affect how

teachers construct and critique arguments. This study explores how pre-service teachers view the inductive-deductive distinction in terms of viable arguments.

Morris (2002) showed that pre-service teachers "used a wide variety of criteria to evaluate student's arguments and rarely used logical validity as a criterion for evaluating the arguments" (p. 510). They also held two different views about mathematical justification—one reserved for explanation, the other for proof. When pre-service teachers focused on explaining a result and understanding why a mathematical statement was true, they readily accepted a *generic example* and rejected example-based inductive arguments. A different viewpoint emerged related to their understanding of proving in mathematics. When pre-service teachers applied this criterion, they indicated that multiple examples prove a statement, whereas a single-case key idea inductive argument does not. In other words, a single-case key idea argument was preferred to explain a generalization, whereas multiple examples were regarded as sufficient to prove the generalization. This study will further investigate how pre-service teachers view generic examples and lists of examples in relation to what constitutes a viable argument.

Knuth (2002) interviewed 16 practicing secondary school mathematics teachers, most with solid mathematical backgrounds, about their conception of proof. Knuth's results revealed some striking misconceptions about the nature and role of proof in mathematics: of the 16 teachers interviewed, 4 tested more examples after a proof had been given, 6 thought it would be possible to find contradictory evidence to a proven statement, 10 accepted the proof of a converse as proof of the original statement, and 13 found empirical examples to be the most convincing type of argument. Additionally,

more than one third of the arguments presented to the teachers did not fulfill the criteria for proof but were rated as proofs by the teachers. Regarding teachers' conception of the role of proof in teaching, Knuth (2002) further reported that all 16 teachers in the study saw verification as the main role of proof in school mathematics, whereas none of the teachers mentioned explanation.

Varghese (2009) had similar results in a study of 17 student teachers completing the final semester of a Canadian teacher preparation program. In Varghese's study, only one of the participants mentioned explanation as a role of proof. These studies seem to indicate that both pre-service and in-service teachers have a rather narrow conception of the role of proof in teaching. However, this is possibly a result of specific connotations of the word *proof* and that different results will be obtained by using *viable argument*.

The above studies seem to indicate that current and prospective teachers often see explanations as distinct from mathematical proof, which means that the process of explaining and justifying a mathematical result is different from proving a mathematical result. This suggests a viewpoint that convincing somebody about the truth value of a statement is different from proving the statement. This may stem from connotations of rigor invoked by the term proof, and a belief that proving a statement is an exercise in logic and formal rigor and different from arguing that the statement is true which may take an informal approach. In this sense, proof becomes a separate mathematical subject, a content topic to be taught rather than a mathematical practice permeating all of mathematics. That this view exists is supported by a teacher in Knuth's study who

explains if that she had to drop something from the curriculum, proof would be the first topic to be removed.

The results from research on teachers' conceptions of proof have generally been interpreted as deficiencies in teacher knowledge and a failure to grasp the importance of proof in the mathematics curriculum. However, at least one recent study (Nardi et al., 2012) points to the fact that classroom interactions are complex and that teachers' conceptions should not only be analyzed in terms of accuracy, but also in terms of the multiple considerations and priorities under which they operate, including pedagogical, curricular, professional, and personal concerns. It is possible that teachers who do not mention explanation as a reason to teach proof nevertheless value justification in the classroom, but interpret proof to have a mathematical rigor not necessary for a mathematical justification. This study will provide insight into whether—and why—pre-service teachers focus on explanation or proof when they construct viable arguments and critique the reasoning of others.

Recommendations for Teacher Education

Stylianides and Ball (2008) argue that teachers' pedagogical content knowledge of proof needs to consist of two types of knowledge: (1) knowledge of *logico-linguistic aspects of proof* and (2) knowledge of *situations for proving*, such as knowing when proof and proving can be introduced in the classroom and what type of proving activity goes with different proof tasks. Teachers need to possess both of these if they are to engage their students in the process of proving in the classroom. The first of these categories is knowledge about the mathematical structure of proofs including definitions

and axioms, logical validity, and modes of argumentation such as the distinction between empirical and deductive reasoning. To complement this knowledge, teachers also need to be familiar with different kinds of proving tasks and proving activities, such as tasks that employ refutation through counterexample or proof by exhaustion.

Stylianides and Ball (2008) used transcripts of teaching episodes to discuss different types of proving tasks and proving activities, and they directly tied the knowledge required for teaching proof to situations encountered in the classroom. They noted in one observation that "the teacher needed knowledge about proof that would help her address the issue of whether Betsy's conjecture could be verified and, if so, how" (p.327). This means that teachers need to have a sound knowledge of how to extend incomplete student arguments into viable arguments and be able to recognize what features of an incomplete or flawed argument can be used to produce a better argument. This study will explore whether pre-service teachers possess the necessary skills to extend and improve student arguments.

Conclusion

This chapter provided an overview of current research regarding the nature of proof, the role of proof and explanation in the K-12 curriculum, and how in-service and pre-service teachers interpret, construct, and evaluate proof tasks. Where possible, the discussion has been extended to include argumentation as an alternative to the classical concept of proof. The following chapter describes the research methodology that will be

used to extend this body of research to include pre-service teachers' interpretation, construction, and evaluation of viable arguments.

CHAPTER 3

METHODOLOGY

Introduction

This chapter begins with a restatement of the research questions as the foundation for the choice of research design and methodology. This is followed by a description of the research design including details about the setting and the participants involved in the research, a discussion of methods for data collection and data analysis, and attention to issues of reliability and validity.

Research Questions

This study focuses on a sequence of related research questions, all with respect to pre-service secondary teachers' understanding of viable arguments and their strategies as to how they "construct viable arguments and critique the reasoning of others" as laid out in the Common Core Standards for Mathematical Practice. The study has three specific and related goals: (1) to investigate how pre-service secondary teachers understand and interpret the notion of a "viable argument"; (2) to investigate the strategies employed by pre-service secondary teachers when they are asked to construct viable arguments; and (3) to investigate how pre-service teachers evaluate student arguments attempting to justify mathematical statements. The research questions are as follows:

1. How do pre-service teachers define the notion of "viable argument"? Is their notion consistent with "viable argument" as defined by the Common Core Standards?
2. What strategies do pre-service teachers employ when asked to construct viable arguments?

   2a: Do they build a logical progression of statements to explore the truth of conjectures?

   2b: Do they use examples constructively in their arguments?
3. What characteristics do pre-service teachers assess to critique the reasoning of others?

   3a: Do they distinguish between correct logic and logic that is flawed?

   3b: Do they look for ways to clarify or improve the arguments of others?

   3c: Do they recognize and value constructive uses of examples?

Research Design

A qualitative research design is most appropriate for exploring and understanding students' mathematical thinking (Hoepfl, 1997). In particular, this study attempts to identify and categorize aspects of student thinking that must be brought into the open through questioning and prompting. A case study methodology relying heavily on task-based interviews was used to illuminate the thinking of a selected sample of undergraduate students, allowing for comparison and contrast of their approach to constructing and critiquing viable arguments.

Since a major goal of this research is to uncover mathematical thinking, this study can be classified as exploratory, and a qualitative research methodology using task-based interviews is appropriate (Hoepfl, 1997; Goldin, 2000; Clement, 2000; Golafshani, 2003; Patton, 2002; Shaughnessy, 2004; Corbin & Strauss, 2008). Using clinical interviews as a means of uncovering cognitive structures and participant understanding is a research methodology pioneered by Piaget that later gained traction among developmental psychologists and mathematics and science education researchers. The strength of this methodology is that it allows the researcher to uncover cognitive structures and beliefs that are not as readily apparent in less open-ended research methodologies (Clement, 2000).

In mathematics education, case studies and task-based interviews have been used for a variety of purposes and the methodology has yielded some important and influential results in mathematics education. For instance, Erlwanger's seminal study (1973) investigating a particular student's misconception about fractions had major implications for mathematics education in that it warned about the dangers of teaching procedures without teaching the proper conceptual underpinnings (Shaughnessy, 2004). Erlwanger's study is also important from a methodological point of view: In his study of a single student Erlwanger revealed the powerful evidence that can be acquired through studying a single case. In the research on proof, reasoning, and argumentation, both semi-structured and task-based interviews are often employed to answer questions relating to proof understanding and to reveal cognitive structures relating to proof. For instance, Knuth's two studies about teachers' conceptions about proof (Knuth, 2002a; Knuth

2002b), Raman's (2003) study introducing the concept of a key idea, and Weber and Alcock's study (2004) outlining the difference between syntactic and semantic proof production are all examples of influential studies using interview formats.

In this study, task-based interviews were conducted with a sample of pre-service secondary teachers after they completed the majority of a teacher preparation program closely following the Conference Board of Mathematical Sciences recommendations for the Mathematical Education of Teachers (MET2, 2012). The interview questions and tasks are of three different types: some of the questions ask specifically about the participants' understanding of the term "viable argument"; some tasks require pre-service teachers to construct their own arguments while "thinking aloud"; and some tasks ask them to critique a set of student responses to a given problem. The mathematical items used in the interview and presented to the participants were all taken from the field of number theory, with one exception taken from the field of geometry. An additional construction item from Calculus was included as a challenging task for participants who might easily complete the other construction tasks. The researcher designed an initial set of questions to establish the participants' mathematical background and to ascertain their understanding and interpretation of viable argument as defined in the Common Core State Standards. All tasks or questions that involved constructing arguments or critiquing student responses have been used in previous studies of proof and reasoning.

Other predominant methods used to discover student thinking include classroom observations and collecting written work. These were considered not appropriate for this study since these methods would introduce several confounding factors. First of all, a

classroom observation of an individual necessarily introduces input and feedback from the instructor and interactions with the instructor and other students. This was not desired since the main purpose of the study was to examine individual pre-service teachers' notion of viable argument, how they construct viable arguments, and how they evaluate arguments. Additionally, tasks introduced in the classroom setting are often concerned with challenging content with the intent to teach new material, so the solution process is highly interactive and involves discourse between several students. It was imperative to the research that the tasks be individual, whereas student work in the classroom setting is often collaborative and does not necessarily capture individual student thinking.

The Researcher

This research is conducted as a partial fulfillment of the requirements of a Doctorate of Philosophy in Mathematics with an emphasis on Mathematics Education. Before completing a Master's degree in Mathematics in the United States, I completed all coursework required for Norwegian High School teaching certification in Mathematics and Physics and taught 9^{th} grade in Norway for a semester. Prior to conducting this study, I gained experience in conducting semi-structured interviews: four interviews with pre-service elementary teachers on their understanding of proof in the elementary grades and three interviews as part of a pilot study to inform the design of the research conducted here.

I have not taught any of the three courses, Methods of Proof, Higher Mathematics for Secondary Teachers, or Modern Geometry, that are aligned with the CBMS

recommendations for engaging pre-service teachers in proof, reasoning, and argumentation. I am familiar with the content of the three courses having substituted on several occasions in Methods of Proof and I completed Higher Mathematics for Secondary Teachers and Modern Geometry as a student, although the courses were titled Discrete Mathematics and College Geometry at that time. My teaching experience relevant to proof, reasoning, and argumentation is limited to teaching Linear Algebra and the Mathematics for Elementary Teachers sequence, a sequence that emphasizes mathematical reasoning, albeit in the elementary grades.

Participants and Mathematical Coursework

The participants in the study were pre-service teachers in their third year of the secondary mathematics teacher preparation program at a midsize university in the Mountain West. In order to graduate with a mathematics teaching degree, students must complete the standard calculus sequence: (Calculus I, Calculus II, Multivariate Calculus, and Introduction to Differential Equations). Additionally, all the different secondary mathematics teaching degrees require a suite of courses designed specifically for prospective teachers: Methods of Proof, Higher Mathematics for Secondary Teachers, Modern Geometry, and Mathematical Modeling for Teachers.

At the university where the study took place, the suite of courses required for a mathematics teaching degree are closely aligned with the recommendations expressed by the Conference Board of Mathematical Sciences in their recent publication, The Mathematical Education of Teachers II (CBMS, 2012). In this document the CBMS

argues that experience with reasoning and proof should be integrated across the entire spectrum of undergraduate mathematics. They state that teaching mathematical reasoning "requires a classroom where learners are active participants in developing the mathematics and are constantly required to reflect on their reasoning" (p. 56). They further claim that teachers, like their students, need to have the expertise described in the Common Core State Standards, which includes constructing viable arguments and critiquing the reasoning of others. At the university in this study, the courses designed to address these recommendations are Methods of Proof, Higher Mathematics for Secondary Teachers, and Modern Geometry.

Participants

Students selected for this study were drawn from the Modern Geometry course. All participants had completed the multiple subject requirements listed above except Mathematics Modeling for Teachers. In other words, they were enrolled in the third of four specialized content courses for teachers, and had completed two of the three courses designed to develop fluency in mathematical reasoning, argumentation, and proof. This choice helped ensure that the participants had considerable experience with mathematical proof, reasoning, and argumentation. Furthermore, the participants are near completion of the three courses that align with the Common Core State Standards call to "construct viable arguments and critique the reasoning of others."

The pool of potential subjects included several non-traditional students who have either returned to college or started college after several years of work experience. Only "traditional" college students were considered for this study, in order to minimize

possible influences from outside academic or work experience. This limitation additionally ensured that the participants completed their entire high school education after 2000, when the National Council of Teachers of Mathematics recommended that proof be introduced at all levels of school mathematics.

A total of eight students enrolled in Modern Geometry satisfied the criteria and were asked to participate in the study. One declined to participate, one did not respond, and one failed to respond after initially agreeing to participate. Of the five participants who were selected, four were juniors and one was a senior completing the final semester of the program. Four were mathematics teaching majors, whereas the fifth was majoring in broad-field science education with a minor in mathematics teaching.

Methods of Proof

Methods of Proof serves as an introductory course to mathematical methods and proof. The content in Methods of Proof uses material from a variety of mathematical disciplines such as set theory, analysis, and group theory to highlight the methods and structure of mathematical proof. The emphasis is on logic and the language and syntax of mathematics including quantifiers and negations. Students are introduced to both direct and indirect proofs as well as proof by induction. This is initially done through the use of rules, language, and symbolism of formal logic, and students spend the first part of the class formulating negations to mathematical statements before they are asked to perform basic mathematical proofs on their own. Since the class is essentially an introduction to mathematical proof, the emphasis in the class is on a rigorous approach to proving.

Higher Mathematics for Secondary Teachers

Higher Mathematics for Secondary Teachers gives an introduction to mathematical concepts and processes relevant to secondary school mathematics. The course acts as a bridge and demonstrates how higher level mathematics is relevant in a secondary school setting. The initial part of the course treats high school problems in an advanced setting, whereas in the remainder of the course students are introduced to group theory, discrete mathematics, operations, number theory, divisibility arguments, and induction. The course emphasizes both mathematical proving techniques as well as informal argumentation, particularly as it is used in discrete mathematics.

Since a significant portion of the curriculum is concerned with discrete mathematics and operations on integers, the method of proof by induction is covered extensively in this course. This implies that students completing the class should be familiar with proof by induction as a method to prove statements and recognize this as a useful approach to prove statements over the set of integers. The course also provides experience in how to formalize patterns through inductive arguments.

Modern Geometry

Modern Geometry approaches geometry from a modern and constructivist perspective rather than the classical Euclidean axiomatic approach. The course integrates dynamic technologies such as GeoGebra and Geometers Sketchpad to generate and test examples with the goal of stating conjectures, understanding theorems, and developing a starting point for a mathematical proof. Class time is student-centered rather than lecture-based and class periods are often spent focusing on students' work. Throughout the

course students have rich opportunities to analyze, critique, and suggest improvements to the work of their peers.

Additionally, Modern Geometry actively employs resources supporting the Common Core State Standards. Students are introduced to the eight Standards for Mathematical Practice and the Geometry Standards for different grade levels. They engage with the Illustrative Mathematics Project Web site and throughout the semester are required to analyze and comment on problems posted on the site.

The focus in Modern Geometry is on mathematical argumentation and justification rather than rigorous mathematical proof. This means that the instructor and students value the understanding of geometrical concepts and key ideas that provide the rationale behind a mathematical result. In keeping with the obvious emphasis on the CCSS, the instructor attempts to introduce language associated with the CCSS content and practices, including the phrase "viable argument."

Sample

Eight students enrolled in the Modern Geometry course were approached and asked to participate in the study. The eight candidates were selected based on the criteria that they had completed both Methods of Proof and Higher Mathematics for Secondary Teachers and that they were traditional students in the sense that they were not returning to education after several years in the workforce. Additionally, the course instructors of Modern Geometry and Higher Mathematics for Secondary Teachers were consulted to ensure that the eight students represented a stratified sample of high-achieving, average-

achieving and low-achieving students. The stratification was meant to achieve the best representation possible of "typical" pre-service secondary teachers. Deliberately selecting for variety in the sample based on achievement allows the researcher not only to consider individual responses, but also to compare and contrast data vertically within groups and horizontally across groups. This supports the trustworthiness of the results, particularly given the small sample size. The stratified sample also created greater depth in the subject pool as the researcher sought to uncover whether prior mathematical achievement influences pre-service teachers' ability to recognize and develop viable arguments and to identify viable arguments in the reasoning of others.

Of the eight pre-service teachers approached, one declined to participate and two did not respond to the researcher's requests. Of the remaining five who agreed to participate, one was rated by instructors as high achieving, two were rated average to high achieving, one was rated as average achieving, and one was rated as low to average achieving. Even though the researcher would have liked to interview one or two more students identified as low achieving, it was determined that the five participants satisfied the goal of interviewing a stratified sample in terms of achievement. However, the obtained sample made comparisons within groups impossible, and no such comparisons were attempted. Arrangements were made with each individual participant, and interviews took place approximately halfway through the semester. The participants received a modest gift card as compensation for their participation in the interviews.

Instrument Development

A task-based interview protocol was developed to investigate the research questions through a series of mathematics problems and related questions. The interview protocol was the centerpiece of data collection for this study, so the quality of the instrument is of utmost importance. It is therefore necessary to provide a detailed explanation of how the interview is structured and the kinds of tasks that are included.

The instrument contains three different types of tasks. The initial questions and the first construction task served the purpose of illuminating the participants' perceptions and definitions regarding the phrase "viable argument." The second portion of the instrument addressed the participants' proficiency in constructing viable arguments, by asking them to determine or justify the truth value of mathematical statements or claims. Finally, samples of student responses to similar tasks were presented as a means to investigate how the participants critiqued the reasoning of others. The instrument thus incorporated tasks where the participants had to construct an argument (*construction tasks)* and tasks where they were presented with arguments and asked to critique the reasoning in these arguments *(evaluation tasks).* The format of having students both construct their own arguments and critique the reasoning of others is similar to methodologies employed by Stylianides and Stylianides (2008) and Healy and Hoyles (2000).

The construction and evaluation tasks to be used in the interview were selected based on a set of criteria related to their mathematical value as well as their potential for eliciting student thinking. First, the main focus of the study was to uncover participant

thinking related to the merits and shortcomings of arguments, rather than to determine the truth value of conjectures. This means that false statements, counterexamples, and solution strategies employed when the truth-value of the conjecture is unknown were not the focus of this study, and consequently the mathematical problems and examples used in the research were limited to true statements over infinite sets (i.e., the result pertains to infinitely many cases and cannot be proven by using examples alone). This choice of items avoids proof by exhaustion as a possible solution strategy, making it easier for the researcher to investigate whether interviewees use inductive or deductive strategies in their argument construction. Likewise, using true statements over infinite sets allows the researcher to investigate how interviewees regard empirical evidence in critiquing the arguments of others. For instance, do they regard empirical evidence to be of little value since it is insufficient and mathematically invalid, or do they look for patterns and structures in the empirical evidence that can help generalize and improve the argument?

Selection of Construction Tasks

Considerations of the Population.

Mathematical content for the tasks was for the most part taken from number theory, except for one geometry item and one calculus item. Since the participants had already completed the course in their program that explicitly addresses number theory, using number theory problems as the basis for the interview ensured that the mathematical context of the tasks would be familiar and allowed the pre-service teachers to focus on argumentation rather than struggle with a possibly unfamiliar context. At the same time, to avoid rote recall of previous results, the construction tasks

needed to be tasks that were not explicitly proven in class. To ensure this, the instructors of Methods of Proof and Higher Mathematics for Secondary Teachers were consulted to verify that the construction tasks had not been explicitly covered in class.

In a similar manner, the evaluation tasks were selected with the notion that a proof of a given task would not be immediately familiar to the participants. During a pilot study it was revealed that one of the evaluation tasks had been an exam question in Higher Mathematics for Secondary Teachers. However, it was deemed unlikely that the participants' responses on the exam would significantly color their responses in the interview, for two different reasons. First, the exam was given in a semester previous to this study and so was likely to be far enough removed in time that the students would not immediately recall their answers. Second, the problem was used in the interview as an evaluation task asking the pre-service teachers to evaluate student responses, whereas on the exam the participants were asked to prove the result. Therefore it is unlikely that the participant responses were influenced by memorization of the solution to the exam question, an assumption further supported by the fact that none of the participants in the pilot study referred to their exam solution when completing the interview task. In fact, one of the participants jokingly expressed regret that the exam was prior to the interview and not vice versa.

Considerations of Difficulty Level. Since the qualitative data for this study were based on responses to the posing of mathematical tasks, those tasks had to be clearly stated, well-defined, of an appropriate difficulty level, and at the same time not immediately familiar to the participants. The selection of appropriate tasks was therefore

not an easy matter. A task-based interview must balance the difficulty level of the tasks, such that the participants are challenged but not overwhelmed. Tasks that are too easy reveal little about students' understanding of viable argument. Tasks that are too hard will not yield trustworthy and adequate data.

To ensure that the interview tasks for this study were of the appropriate level, the researcher used two measures. First, the interview tasks were discussed in a seminar with the mathematics education faculty, including the instructors of the three classes addressing proof, reasoning, and argumentation. The general consensus among the three instructors was that the items were of the appropriate difficulty level, and if anything even a little too easy. Second, a pilot study was conducted a semester earlier with three students at nearly the same stage of the teacher education program when the pilot study was conducted as the participants in the study when the study was conducted. The pilot study was invaluable in striking the desired balance in difficulty among the interview tasks. In the pilot study, all three participants completed the first construction task, which yielded information on the pre-service teachers' preferred argument construction method when they readily see an appropriate solution strategy. However, two of the participants were not able to successfully complete the second task, which yielded data on their cognitive strategies when a solution is not immediately obvious. The third participant solved both construction tasks without much problem. Based on these observations, the interview tasks were deemed to be of the appropriate difficulty level, with the caveat that some higher-achieving students might not be challenged by the original problems. As a

result, a more challenging calculus task was selected as a back-up problem for students who might be able to immediately prove the original construction tasks.

Considerations of Reliability. To address issues of trustworthiness, credibility, and dependability, the interview tasks were all taken from existing studies on proof and reasoning where they were previously tested and yielded valuable information about students' mathematical understandings in relation to proof. The source of each task is described in the following paragraphs.

The initial construction task asks students to show that the sum of the angles in a triangle is 180°; this task was employed by Knuth (2002 b) in his study of 17 experienced secondary school teachers' conception of proofs in the context of secondary school mathematics. In one of his items, teachers examined an empirically-based argument where the angles in the triangle are cut out and put together to form a straight line; five of them rated this as a proof. Two of the tasks were taken from Martin and Harel's study of *proof frames* held by future elementary teachers (1989). This study of over 100 pre-service elementary teachers indicated that many of them employ and value inductive means of verification: 80% of the subjects gave a high validity rating and over 50% gave a very high validity rating to at least one inductive argument. Furthermore, the study indicated that the transition from an inductive proof frame to a deductive proof frame is not absolute, as over a third of the students valued both inductive and deductive verification methods.

Another interview task, used both as a construction task and an evaluation task, was adapted from Stylianides and Stylianides (2009) in their study of 39 prospective

elementary teachers' abilities to construct and evaluate proofs. In this study, the particular task, to prove that the sum of two consecutive odd numbers is a multiple of four, was given on an exam. The prospective teachers were specifically asked to avoid an algebraic proof of the statement since this would not be appropriate for elementary school students. It was found that some participants provided empirical arguments even though they knew that these arguments were mathematically invalid. One reason for this was that they were unable to produce a valid deductive argument.

The more challenging calculus item used as a back-up problem was used by Raman (2003) in her study of views about proof in college calculus students and their teachers, i.e. mathematics graduate students and faculty members. In this study, Raman argues that there is an essentially private and an essentially public aspect of proofs and that key ideas; ideas that both explain and convince, provide the link between the private and the public sphere. Raman found that mathematicians essentially see proof as being about uncovering the key idea, whereas students do not.

Selection of Evaluation Tasks

As a framework for selection of the evaluation tasks, the researcher applied an argumentation hierarchy that blends criteria from two different argument classifications developed in previous research; Stylianides and Stylianides (2009) and Balacheff (1988). In their study of 39 pre-service elementary teachers' abilities to construct and evaluate proofs, Stylianides and Stylianides classified the participant responses into five different categories. These categories are:

- *Proof*

- *Valid general argument but not a proof*
- *Unsuccessful attempt for a valid general argument* (i.e. invalid or unfinished general argument)
- *Empirical argument*
- *Non-genuine argument* (i.e. response that showed minimal engagement, irrelevant response, or response that was potentially relevant but the relevance was not made evident by the solver)

The authors argue that these categories are hierarchical and given in the order of decreasing sophistication from a mathematical standpoint. As discussed earlier, Balacheff (1988) classifies empirical arguments into three different categories of increasing sophistication: *naïve empiricism,* where the argument consists of a list of examples: *crucial experiment,* the testing of special cases; and *generic example,* one specific example that contains features from which a proof can be constructed.

Both of these models contain features that are important to constructing viable arguments and critiquing the reasoning of others, but they also contain extraneous characteristics that are not useful to this study. The model offered by Stylianides and Stylianides (2009) is originally constructed with respect to proof and therefore emphasizes mathematical validity. Consequently, a logical but invalid attempt at a generalization is ranked as more sophisticated than empirical evidence. However, as Balacheff (1988) has shown, there is considerable difference in merit in different types of empirical evidence, and it is important to build this into any hierarchy of argumentation. One problem with Balacheff's categorization is that it is often difficult to discern which

category a written response to a conjecture falls into (Varghese, 2009). For instance, a response that gives three sequenced examples to the conjecture that an odd number plus an odd number is even might be categorized as naïve empiricism, as the author of the argument only provides a few examples. However, it is also possible that the author recognizes an inductive pattern but simply fails to explain this, in which case the argument is not naïve at all.

The argumentation hierarchy used in this research study attempts to address these two concerns by minimizing the emphasis on mathematical validity and by deliberately avoiding attempts to intuit the intention of the argument's author. Instead, the focus is on whether the argument contains generalizable features or not. To categorize different levels of argument from this perspective, the researcher regards mathematical argumentation as having five different levels. Following Stylianides and Stylianides (2009), the levels are viewed as hierarchical with each level representing different levels of sophistication in mathematical argumentation. Following Balacheff (1988), empirical evidence is judged on its merit and inherent potential rather than on its mathematical validity. The five different levels, in decreasing order of sophistication, are *formal proof, generalized argument, generalizable argument, non-generalizable argument,* and *irrelevant argument*. The hierarchy can be visualized in Figure 1. The shape of a pyramid has been adopted to reflect the increasing level of sophistication.

In the context of this argumentation hierarchy, an argument is considered a *proof* if it employs formal deductive logic in that it uses definitions, prior results, and deductive

reasoning to rigorously establish the truth of a mathematical statement in a manner consistent with its standard use in college level mathematics.

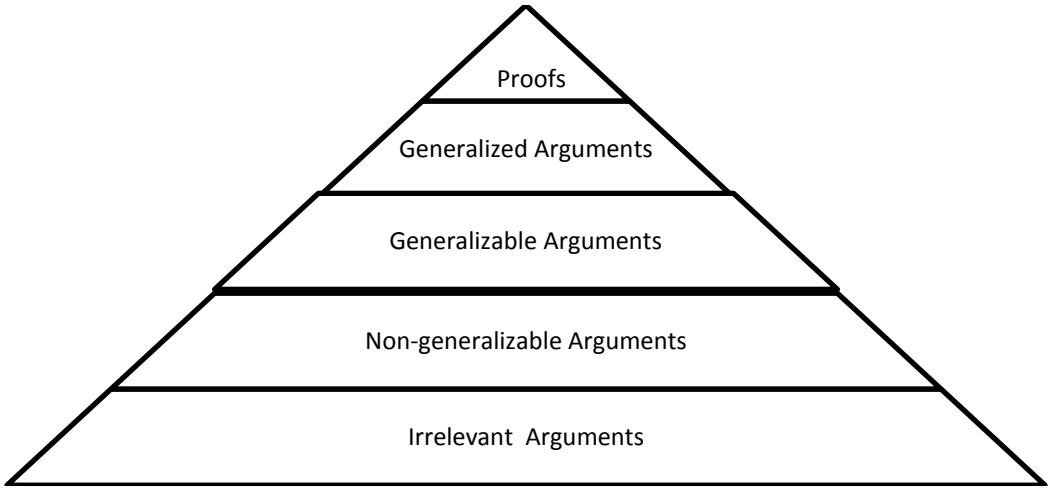


Figure 1: Argumentation Framework

A *generalized argument* is an argument that lacks the formal rigor of a mathematical proof, but contains and explains the important ideas and characteristics from which a formal proof can be written. A generalized argument can be regarded as a pre-formal proof (Blum & Kirsch, 1991), which can be defined as "a chain of correct, but *not formally represented conclusions* which refers to valid, *non-formal premises"* (p. 187, italics in original) or as a blueprint or a prototype for a valid proof (Stylianides & Stylianides, 2009). A generalized argument can be inductive, empirical, or in pictorial form, but all generalized arguments share two features. The first is that the claim of the argument and its generality can be deduced directly from the argument at hand, and, if formalized, the arguments would correspond to a valid mathematical proof. The second universal feature is that the argument explicitly points out the generalizing feature of the

argument and explains how the argument is independent of the concrete example at hand. Examples of generalized arguments include those that Balacheff labels *generic examples*, where the structure of a valid proof is given in one specific example. Yopp (2009) also includes inductive arguments that contain and explain inductive patterns without the formal rigor of a proof by induction. Pictorial arguments constitute a third type of generalized argument where the specific case mentioned intuitively holds true for all possible cases. Blum and Kirsche (1991) refer to these as *geometric-intuitive proofs;* e.g., using an array to visualize the commutative property for multiplication. Using dynamic software to explain mathematical properties also falls into this category.

A *generalizable argument* is an argument that consists of one or more examples containing features or patterns that can be extended to a general argument or to a formal proof, but where this feature or pattern is not explicitly addressed. A generalizable argument differs from a generalized argument in that the generalizing features are not explicitly addressed or immediately obvious in the argument. Some arguments previously classified as naïve empiricism actually contain features or patterns that could be extended to an inductive proof.

A *non-generalizable argument* is an argument that consists of one or more examples, but where key features or patterns that support a generalization are not present and the argument cannot be extended to a formal proof. Among non-generalizable arguments are *crucial experiments* where special cases are tested and the failure to find a counterexample is taken as proof of the statement's truth. Also in this category is creation of a list of examples where it is not possible to extend the argument inductively. These

types of arguments represent Balacheff's naïve empiricism (1988), and the belief that empirical evidence is enough to prove a statement is referred to as the *empirical proof scheme* (Harel & Sowder, 2009).

Finally, an *irrelevant argument* is an argument that has no bearing on the problem or where the bearing on the original statement is not clear.

Description of Instrument and Tasks

Uncovering Pre-Service Teachers' Conceptions of Argumentation

The first part of the interview is devoted to questions attempting to reveal how pre-service teachers understand and interpret the concepts of proof, reasoning, and argumentation in mathematics with particular attention to the notion of "viable argument." These questions are modeled after questions employed by Knuth in similar studies of 16 in-service secondary mathematics teachers (2002 a) and 17 experienced secondary teachers (2002 b). As shown below, a series of prompts were used to draw out each participant's unique understanding of argumentation as they considered (1) the meaning of viable argument in their own mathematical work; (2) the meaning and role of viable argument in school mathematics; and (3) the expectations for viable argument set out in the Common Core State Standards. The prompts used to elicit these understanding included:

- What does the notion of a viable argument mean to you?
    - What does it mean to provide a viable argument?
    - What characteristics qualify an argument as viable?

  - What is the difference between a proof and a viable argument, if any?

- When should students encounter viable arguments?
  - What would you consider a viable argument in elementary school mathematics?
  - What would you consider a viable argument in secondary school mathematics?
- (Upon reading the relevant Common Core Standard for Mathematical Practice) How do you interpret the definition of "viable argument" provided by the writers of the CCSS?
  - Does reading this definition change any of your previous thoughts about viable arguments?

It is important to recognize that since the interview was semi-structured, the researcher made decisions about which questions to ask during the course of the interview.

Interview Items for Construction

During the construction portion of the interview, the pre-service teachers were asked to argue the truth value of three conjectures. The first construction item is the following:

Construction Item 1: Describe how you would construct an argument to support the following true statement: **The sum of the angles in a triangle is 180°.**

Figure 2: Construction Item 1 from Knuth (2002b)

The researcher's (perhaps incorrect) assumption was that pre-service teachers at this stage in their preparation would have knowledge of two or three different methods to justify the statement. After completing this task, the pre-service teachers would be asked if they knew any other methods of answering the question and which of these methods, in their opinion, constitutes a viable argument. The purpose of Construction Item 1 is therefore largely to serve as an introduction to a discussion of viable arguments. However, a secondary purpose is to also investigate if pre-service teachers tend to choose to formally prove the conjecture or if they choose a more pedagogically practical, if less elegant, argument. In the pilot study, one participant proved the conjecture whereas the other two provided different arguments explaining why the argument is true.

The second construction item is adapted from the Stylianides and Stylianides study of how 39 prospective elementary teachers construct and evaluate proofs:

> **Construction Item 2:** As a participant in a teacher workshop focusing on reasoning and argumentation in mathematics, you are asked to provide a viable argument to answer the following conjecture:
> **Is the sum of two consecutive odd numbers divisible by four?**
> How do you respond?

Figure 3: Construction Item 2 Adapted from Stylianides & Stylianides (2009)

When Stylianides and Stylianides administered this problem to prospective elementary teachers, approximately 25% (9 out of 39) gave a solution that was coded as an empirical argument. From this they concluded that a significant portion of the prospective elementary teachers in their study seem to hold the belief that empirical evidence can be used to prove a statement like this. It is important to note that one of the goals of this study was to provide the pre-service elementary teachers with ways of

proving without algebraic representations. The participants were therefore instructed to not give algebraic answers, which may have affected the result.

For the third construction item, pre-service teachers in this study were asked to construct a viable argument to justify the transitivity of divisibility. This problem was used by Martin and Harel (1989) in their study about proof frames in 101 pre-service elementary teachers, where the problem was chosen for its likelihood to be a generalization unfamiliar to the participants.

Construction Item 3: Provide a viable argument <u>for or against</u> the following conjecture: **If a divides b and b divides c, then a divides c.**

Figure 4: Construction Item 3 from Martin & Harel (1989).

This item was chosen on the basis that it contains a key idea—*a* must be a common factor of *b* and *c*—an idea that can act as a bridge between an empirical argument and a more formal mathematical proof (Raman, 2003; Yopp, 2009). Of importance in the selection of this interview item was the fact that the property was not directly covered in the Methods of Proof or Higher Mathematics for Secondary Teachers. This is important since research has indicated that students tend to do better on proof construction tasks that are familiar to the participants (Healy & Hoyles, 2000). At the same time, the definition of divisibility was covered in at least one of the two classes, ensuring that the problem was at an appropriate difficulty level for the participants.

Interview Items for Evaluation

During the interview, the pre-service teachers were presented with two evaluation tasks asking them to critique the reasoning in students work generated by the researcher. The first evaluation task was given immediately after the participants had constructed a viable argument to complete Construction Task 2, showing that the sum of two consecutive odd numbers is divisible by four. The task displays five different researcher-generated responses to the same conjecture:

Is the sum of two consecutive odd numbers divisible by four?

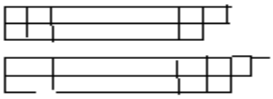
Pat's response: 1+3=4, 5+7=12
7+9=16 11+13=24
53+55=108 123+125=248
We see that all the answers are divisible by four so the sum of two consecutive odd numbers is divisible by four

Chris' response: 1+3=4
3+5=8
5+7=12
We see that the sum of two consecutive odd numbers is divisible by four.

Alex's response: Consider two consecutive odd numbers. One of these numbers must be one more than a multiple of four and the other must be one less than a multiple of four. For example, 25 is one more than a multiple of four and 27 is one less than a multiple of four. So, when you add them together, you get a multiple of four. 25+27=52, which is a multiple of four.

Taylor's response: If you add two odd numbers the answer is always even. When you add two even numbers, the answer is always a multiple of four. So, when you add two consecutive odd numbers the answer is always a multiple of four.

Ryan's response:



When these two odd numbers are added together, we see that the one box left over in the larger odd number will fill in the missing box in the smaller odd number creating a new box with four rows and no boxes left over. Since there are four rows, the number is a multiple of four. This means that the sum of two consecutive odd numbers must be a multiple of four.

Figure 5 : Researcher Generated Responses for Evaluation Task 1.

The five responses include a general argument lacking the formal rigor of a mathematical proof in Ryan's response. This response was constructed by the researcher, but is similar to a student response coded as a proof by Stylianides and Stylianides (2009) in their study of proof construction and evaluation among pre-service elementary teachers. Furthermore, there are two generalizable arguments, since Chris' response can be extended inductively and Alex's response uses a generic example that can be extended to a valid proof. Pat's response is an example of a non-generalizable argument, since the argument does not contain features that are generalizable. Lastly, Taylor's response is an example of an argument that is logically flawed; it is not apparent how the conclusion follows from the two given premises, and one of the premises is false since the sum of two even numbers is not necessarily divisible by four.

The second evaluation task consists of four different researcher-generated arguments supporting the rule for divisibility by three. The rule is stated in figure 6, followed by the four research-generated responses. The four different responses include one proof, albeit only for 3-digit numbers, in Gus' response, a generalizable argument in Marina's response, a non-generalizable argument in Reece's response, and a logically

flawed or incorrect argument in Anna's response. This problem was employed in Martin and Harel's (1989) study of pre-service elementary teachers as well as in Knuth's (2002b) study of in-service secondary teachers. Yopp (2009) has used the divisibility of three rule as an example of a mathematical result employing a key idea that can serve as a bridge between empirical evidence and deductive justification.

A number is divisible by 3 if the sum of the digits is divisible by 3.
Reece's response:

Number	Sum of Digits	Number divisible by 3?	Sum of digits divisible by 3?
12	3	Yes	Yes
13	4	No	No
360	9	Yes	Yes
361	10	No	No
137,541	21	Yes	Yes
157,541	23	No	No

We see that the rule is true.

Gus' response:
Let a be any three digit number, with digits x,y, and z. By the place-value concept, a=100x+10y+z. This equality can be written a=(99x+x)+(9y+y)+z. By the commutative and associative properties we get a=(99x+9y)+(x+y+z). Notice that the expression 99x+9y is always divisible by 9, and therefore also by 3. Now, if the second expression, which is the sum of the number's digits, is divisible by 3, then we get that the number itself is divisible by 3.

Anna's response:
Let a be any whole number such that the sum of its digits is divisible by 3. Assuming its digits are x,y, and z, then a=xyz. Since x+y+z is divisible by three, also xyz is divisible by three. Therefore a is divisible by 3

Marina's response:
Consider 756. This number can be represented as follows: 756= 7x100+5x10+6x1. This can be rewritten 756= (7x99+7)+ (5x9+5)+6. By the commutative and associative properties, we get 756=(7x99+5x9)+(7+5+6). Notice that the expression 7x99+5x9 is always divisible by 9, and therefore also by 3. Now if the second expression which is the sum the number's digits, is divisible by 3, then we get that the number itself is divisible by 3.

Figure 6: Researcher Generated Responses to Evaluation Task 2 (Adapted from Martin & Harel, 1989; Knuth, 2002b)

In each of the two evaluation tasks, pre-service teachers were asked to decide which responses are viable arguments and what specific features of these arguments make them viable (or are missing to make them non-viable). They were also asked if the arguments could be improved and how. Finally, if needed, they were prompted to compare and contrast the arguments. The researcher used these responses to investigate which features of an argument the pre-service secondary teachers appeared to value, whether they were able to recognize key ideas, and whether they recognized and attempted to generalize arguments that include generalizing features.

Previous studies have included "leading" prompts about whether the stated argument proves the conjecture, or whether it covers all cases (Martin & Harel, 1989; Healy & Hoyles, 2000). This line of inquiry was avoided for fear it might influence pre-service teachers' evaluation of the argument. Keeping the wording neutral and asking only for an evaluation of the argument helped to avoid steering the responses in one direction or another. Asking if an argument can be improved served to measure whether the participants could distinguish between generalizable and non-generalizable arguments. This prompt also assessed whether the pre-service teachers recognized key ideas and inductive patterns and if they could extend these arguments towards a more formal proof.

Pilot Study

A pilot study was conducted in Fall 2012 with three students enrolled in Higher Mathematics for Secondary Teachers. The pilot study provided valuable data to refine the

design and implementation of the interview protocol, including adaptations to the difficulty level of each task. The students selected for the pilot had already completed Methods of Proof and Modern Geometry, and so had similar exposure to two courses that address reasoning and proof. One of the participants was classified as high achieving by the faculty member teaching the course, one was classified as average to high achieving, and the third was classified as average to low achieving, demonstrating the same breadth of mathematical achievement as the sample used in this study. All the participants were cooperative, seemed eager to participate in the research, and displayed interest in the subject matter.

All task-based interviews in the pilot study were conducted by the researcher. Two of the participants were each interviewed in one session lasting approximately 75 minutes, whereas the third participant was interviewed over two sessions lasting approximately 45 minutes and approximately 30 minutes, respectively. The break between the two-part interview sessions was two days. Based on responses, the researcher determined that both formats are acceptable for completing the interview. During the interviews, the researcher referred to a printed protocol that included the interview tasks and suggested prompts to further probe students' thinking. The researcher also supplemented with additional questions not in the printed protocol as necessary, an acceptable practice in a semi-structured interview situation.

The pilot study took an abbreviated approach to examining pre-service teachers' interpretation of "viable argument." The three participants were first asked what that phrase meant to them. After responding, they were asked to read the Common Core

Standard for Mathematical Practice relating to viable arguments, then asked if they would like to adjust their initial response. A more extensive protocol was used in the current study to thoroughly probe understanding and interpretation of "viable argument."

The four construction items were each read aloud to the participants and repeated and explained in more detail if needed. The participants were given pen and paper to work on each construction task, and their written responses were collected for analysis at the end of the interview session. The interviews were also video recorded and transcribed for analysis. The two evaluation tasks were presented to the participants in no particular order, to minimize bias from any perceived sequencing of the examples. Each evaluation task was provided by the researcher on a separate sheet of paper, and each student response had ample space for the participants to make written comments. Their comments were collected at the end of the interview session.

Results from the Pilot Study

Interpretation Tasks. The three participants in the pilot study had very similar responses in their understanding of what is meant by the term "viable argument." All three participants focused on justification and explaining work as the main purpose of a viable argument. However, the responses were somewhat shallow, and the prompts used in the pilot study failed to capture some of the intent behind the researcher's questions. Consequently, additional questions based on Knuth's research into proof conceptions were added to the final interview protocol.

Evaluation Tasks. When asked to evaluate students' arguments and decide whether these arguments were viable, the participants' criteria for what can be judged as viable differed significantly. For example, the first participant equated a viable argument with a proof; that is, a mathematically valid, rigorous, and formal argument. Consequently this participant considered very few of the student arguments viable. This is interesting seen in the light that this participant explicitly indicated a difference between viable arguments and proof earlier in the interview: "I mean it is not a solid proof but it is a viable argument, because I was able to take any triangle, cut off their edges and make a 180 degree line." The second participant rated all of the student arguments as viable and seemed to pay more attention to the conclusion of the argument—the truth of the statement—than to the structure and content of the argument. This participant readily accepted a flawed argument as viable, because the argument reached the correct conclusion. The third participant invented a scale of viability from 1 to 10 and rated the arguments accordingly.

This spectrum of responses indicates that not only did the participants understand the term "viable argument" differently, they also were not necessarily consistent in what they rated as a viable argument. These results further supported the researcher's decision to strengthen the initial portion of the interview related to understanding and interpretation. Additionally, it was the researcher's original intent to avoid direct reference to proof in the interviews, but results from the pilot study indicated that investigating the difference in the participants' conceptions of "proof" versus "viable

argument" could prove fruitful. Consequently some of the participants in the interviews were asked directly about the difference between "proof" and "viable argument."

Construction Tasks. None of the participants used empirical evidence in their construction of viable arguments. All three were able to produce an algebraic proof for the second construction item by defining an odd number as 2n+1 and then representing the sum of two consecutive numbers as (2n+1) + (2n+3) = 4(n+1), which is divisible by four. However, only one of the participants was able to construct a proof for the second construction item about the transitivity of divisibility. Interestingly, this participant immediately saw the key idea of using common factors and was able to produce a valid algebraic proof with little difficulty. The other two participants attempted to manipulate symbols to produce a proof, but were not successful in doing so.

The analysis of the pilot study subjects' failed attempts to construct viable arguments suggests they may have been limited by a belief that a viable argument has to be algebraic, or that they equated a viable argument with a more rigorous proof. This is further supported in that none of these participants tried to use empirical evidence to assist with the process of developing a "viable argument." In fact, the only time empirical evidence was used was in searching for a counterexample. Absent a counterexample, the participants hypothesized that the statement was true, but they did not use examples to try to visualize why the statement needed to be true. One participant mentioned that there needed to be a common factor, but was not able to support this idea or translate it to an example or an algebraic proof. This may indicate a struggle with recognizing and developing key ideas into a more "viable argument."

Inconsistencies. Intriguing contradictions, or at least inconsistencies, arose in the participants' responses. For instance, the third participant mentioned generality as a trait of a viable argument. Nevertheless, this participant rated a non-generalizable argument using randomly chosen large numbers as being better than a generalizable argument using three well-chosen initial values. The first participant explicitly differentiated between viable arguments and proofs when stating a definition of viable arguments, but made no distinction between the two when evaluating student arguments. The second participant appeared to value justification in a viable argument. Yet instead of focusing on how well the students backed up their arguments, this participant focused on the correctness of the conclusion. These and other findings helped in restructuring and refining the interview protocols.

Data Analysis

This section begins with a brief statement about how previous research results interact with data analysis in this study. It then describes how the results from the pilot study informed data analysis for the current study. The data analysis plan for this research study is then presented, including a detailed description of how the data were analyzed for the three main types of tasks in the task-based interviews.

The study is unique in its focus on pre-service teachers, argumentation, and the Common Core Standards for Mathematics. However, the methodology is not designed to build new theory, but to add to the research knowledge base by illuminating pre-service teachers' understanding of argumentation against the backdrop of existing theory on

proof. The role of proof and reasoning in mathematics has been widely studied, resulting in the development of well-known proof schemes and categorizations by researchers in this area. In this study, it is expected that many of the student responses related to argumentation will be aligned with results reported in previous research on proof.

How the Pilot Study Informed Data Analysis

The results from the pilot study indicate rich variation in pre-service teachers' understanding of the term "viable argument" as well as their approach to argument construction and evaluation. The differences and inconsistencies within individual responses highlighted the need for a layered approach to analyze pre-service teachers' interpretation, construction, and evaluation of viable arguments. Participant responses were examined from a variety of perspectives, including a vertical approach treating each participant as a unique subject of study and a horizontal approach to look for themes that were consistent across all participants.

Although the researcher did not assume that findings would be generalizable to larger populations of pre-service teachers with different mathematics backgrounds, it was anticipated that useful information would be generated to support the preparation of teachers who need to embrace the Common Core State Standards' perspective on viable arguments. For example, it was hoped that the data might reveal common inconsistencies between pre-service teachers' views on viable arguments and their actual use or critique of arguments, or common misconceptions about the use of examples in building a convincing argument. Such results could be used not only to improve pre-service

teachers' own use of viable arguments, but their ability to recognize and diagnose barriers and challenges among their own students.

Initial Analysis

Initial impressions were developed while the researcher conducted the face-to-face interviews. Following completion of all the interviews, the researcher watched the video recording of each subject's interview, which resulted in confirmation or modification of initial impressions. Each interview transcript was then transcribed by the researcher, which allowed for a third review of the live interview. The transcript was broken up into three sections: the answers to the initial task of defining viable argument; responses from the construction and assessment items from the first interview session; and responses to the construction and assessment items from the second interview session.

Each section of an interview was first coded for key phrases and recurring themes based solely on the subject's statements. A second coding then took place using rubrics developed prior to data collection and based on results from the literature on proof, reasoning, and argumentation. The rubrics used are outlined in greater detail later in this chapter. With the initial coding, the researcher wanted to ensure that potentially important ideas would be identified without the bias created by working from pre-established codes. Using a pre-established rubric in the second phase of the coding allowed the researcher to measure findings from this study against prior results from the already existing literature on proof, reasoning, and argumentation. The researcher also acknowledges a detailed

knowledge of the results from the literature and the effect that this by necessity has on the codes and themes used in this study.

A "participant profile" was created for each of the five subjects. Along with providing demographic data about undergraduate experience and overall achievement, the researcher constructed a description of each subject's demeanor during the interview and general impressions regarding each subject's mathematical proficiencies and confidence in relation to proof, reasoning, and argumentation. This complements a more analytic description of each subject's responses to the interpretation, construction, and evaluation portions of the interview and how those responses are situated in the rubrics and coding schemes built from the research literature.

After describing each subject in detail, the researcher adopted a horizontal perspective to examine trends that span all or some of the five cases. This was accomplished by looking for patterns using the initial coding scheme as well as the research-based criteria and external coding schemes. Finally, the researcher sought to identify and report unique or interesting responses. In situations where qualitative data is used to answer questions, a singular or "outlier" response may be as informative as a pattern of responses.

Analysis of Responses to Interpretation Tasks

Data from the first portion of the interview was first coded independently before it was filtered through a rubric based on results from Knuth's studies (2002a, 2002b). Questions in this portion are related to pre-service teachers' understanding and interpretation of the meaning of "viable argument." While the rubric shown in Figure 7

does not produce a definitive analysis, it serves to sort responses into a scheme that has proven useful in studies of proof and reasoning. Knuth's work took place prior to development of the Common Core State Standards, which explicitly state characteristics of a viable argument. These features were incorporated into the rubric in an effort to identify whether pre-service teachers use commonly held language or key words when expressing their understanding of "viable argument."

Table 1: Rubric for Analysis of Definition Items

Does the subject refer to:	**Yes**	**No**	**Codes and Key Phrases**
Explanation?			
Ability to convince?			
Verification?			
Justification?			
Communication and explaining thinking?			
Using valid methods?			
Using logical progression?			
Generality?			

Analysis of pre-service teachers' responses to the interpretation tasks used the same analytical-inductive methodology adopted by Knuth (2002a, 2002b). The rubric in Figure 7 provides a set of external codes grounded in prior research and in the language of the Common Core. In this method the initial *external codes*, developed prior to conducting the interview, are supplemented by a set of *internal codes* as new themes emerge from the data. As data analysis proceeds, the coding scheme can be restructured

based on the applicability of the original external codes and on the emergence of more relevant internal codes. In this way, the coding scheme can be built on prior research, but adapted to suit the new context. The researcher found this approach an effective way to apply pre-existing knowledge about proof and reasoning while watching for evidence of ideas and behaviors that did not fit the coding scheme.

From the data produced by tasks and questions related to interpreting the meaning of "viable argument," the researcher began to build an individual profile for each pre-service teacher based on his or her self-reported views and understandings about argumentation. These profiles would later be extended by researcher observations and the participants' verbal evidence and written statements as they constructed and evaluated arguments in later portions of the interview. Through this cyclic approach to data analysis, the researcher sought to characterize the thinking of the individual participants but also to identify commonalities across the group of participants. The data were also analyzed for consistencies within a single individual's responses and across individuals.

Analysis of Responses to Construction Tasks

Table 2: Rubric for Analysis of Construction Items

Code	Coding Criteria	Example: Construction Item 2	Example: Construction Item 3
P	The argument is recognized as a mathematical proof according to the canonical definition of a proof.	Two consecutive odd numbers are given by 2n+1 and 2n+3 where n is a natural number. Then (2n+1)+(2n+3)=4n+4=4(n+1), which is divisible by four.	If a|b and b|c, then b=an and c=bm for some natural numbers n and m. Then we have: c=bm=(an)m=a(nm), nm is a natural number so a|c

Table 2 (cont): Rubric for Analysis of Construction Items

Code	Coding Criteria	Example: Construction Item 2	Example: Construction Item 3
Gen	Valid argument lacking the formal rigor of proof. The generality of the argument is explicitly addressed. The argument contains qualifiers and explains how the argument can be extended.	The first pair of consecutive odd numbers is 1 and 3. 1+3=4, which is divisible by four. To get the next pair of odd numbers you increase the 1 and the 3 by 2 each, so the sum will increase by 4. Therefore the sum of the next pair of consecutive odd numbers will be divisible by four. Since every time you move from one pair of consecutive odd numbers to the next pair you are adding 2 to each of the numbers in the first pair, the sum will increase by four. This pattern will always hold.	3|6 since 6=3x2. 6|24 since 24=4x6. We have that 3|24 since 24=4x6=4x(2x3). So, since 3 is a factor of 6 which again is a factor of 24, the 3 is a factor of 24.
Gble	Argument that contains features or objects that can be used to obtain a proof, but that does not explicitly address the general nature of the argument. The argument does not contain qualifiers that explain how the argument can be extended.	1+3=4 3+5=8 5+7=12 We see that the sum of two consecutive odd numbers is divisible by four.	3|6 since 6=3x2. 6|24 since 24=4*6. We have that 3|24 since 24=4x6.
NG	Purely empirical argument that cannot be generalized. The argument rests on one or more inductive warrants.	1+3=4 21+23=44 57+59 =116 233+235=468 We see that the sum of two consecutive odd numbers is divisible by four.	3|6, 6|24 and 3|24 25|125, 125|625 and 25|625
IA	Arguments that do not address or are irrelevant to the given conjecture.	21+25=46 is not divisible by four, so the statement is false.	(a\b)x(b\c)=a\c so the statement is true.

The responses to the construction items were analyzed using the five-level argumentation framework presented earlier. Each response was either coded as a proof (P), as a generalized argument (Gen), as a generalizable argument (Gble), as a non-generalizable argument (NG), or as an irrelevant argument (IA). This rubric is a slight revision of the rubric employed by Stylianides and Stylianides (2008) in their study of proof construction and evaluation among prospective elementary teachers. As in that study, the levels are hierarchical and ordered in decreasing levels of sophistication. This study's use of the rubric differed from that of Stylianides and Stylianides (2008) in that the categories were not coded or ranked in terms of validity, but rather in terms of merit relevant to argumentation. The structure of an argument and if the argument included generalizing features was deemed more important than the validity of the argument. As a consequence, mathematically incomplete arguments, which are not highly valued in the Stylianides & Stylianides study, might be regarded here as general or generalizable based on the structure of the argument.

Table 2 provides more detailed coding criteria, along with fictional example arguments from the construction section of the task-based interview that meet the criteria for each code. It is important to note that during the actual interviews, the study subjects supplemented their written arguments with narration and verbalization of ideas. Therefore, the researcher had to consider both written work and verbal statements when coding the arguments.

The pilot study revealed that pre-service teachers may at times construct some form of an argument while recognizing that it is not complete or that it is not fully

convincing. To further explore this possibility, upon completing each construction task the participants in this study were asked directly whether the argument they constructed was viable. Further prompts led them to elaborate on what features made the argument viable, or what features were missing for the argument to be considered viable. This line of questioning helped the researcher determine what types of arguments the pre-service teachers considered viable and how this was related to generalizations.

Analysis of Responses to Evaluation Tasks

In two different evaluation items, the pre-service teachers were presented with a variety of mathematical arguments for the same task and asked to decide if they considered each argument to be viable or not. For arguments that they identified as viable, the participants were asked what characteristics or features make the argument viable. For arguments they deemed not viable, they were asked if the argument could be made viable and what features of a viable argument were missing.

The pre-service teachers were also asked to compare and contrast the different responses provided for each task. Particular attention was given to how they rated empirical arguments against more formal arguments and if they were able to highlight differences in quality between empirical arguments. As with the construction items, a predetermined framework of external codes was used to initiate data analysis for student responses on the evaluation items. This coding framework was developed by the researcher. Rather than representing a hierarchy of levels, this coding scheme was constructed to capture a range of potential responses that might help to answer the research questions. Participants' responses were analyzed and coded based on the rubric.

Table 3: Rubric for Analysis of Evaluation Items

Response Criteria: does the participant…	Yes	No	Codes
Rate the argument as viable?			
Note that the argument can be extended or improved to become viable?			
Address whether the argument is empirical or deductive?			
Address how empirical evidence is used?			
Note features in the argument that can be used to generalize the argument?			
Discuss other issues of generality?			
Discuss logical aspects of the argument?			
Highlight key ideas?			
Discuss the presentation or the representation of the argument?			
Offer suggestions for improvement based on the content of the argument?			
Offer suggestions for improvement based on the presentation of the argument?			

Issues of Validity and Reliability

Measures of validity and reliability for qualitative research do not apply in the traditional quantitative sense. This is not unexpected since qualitative methods are subjective and qualitative data is subject to interpretation – neither of these conditions is

typical of quantitative research. Lincoln and Guba (1985) proposed a framework for validating qualitative research that includes attention to *credibility*, *transferability*, *dependability*, and *confirmability*. There is some overlap among these terms and with other popular terms such as *trustworthiness* and *rigor*. A complete discussion of how these concepts relate to reliability and validity is beyond the scope of this chapter. Instead, this section will focus on ways that the study will maintain high standards of quality that add to its trustworthiness.

First, multiple methods of data collection for this study will help to maintain credibility and accuracy of the data. The participants' responses will be interpreted in three ways: "live" during the interview sessions, as a visual record using digital video, and in verbal form through typed transcripts. The combination of hearing participants' thoughts firsthand, analyzing their words through transcripts, and reviewing their statements and gestures through video will ensure accurate interpretation of the data. Credibility will be further enhanced by creating detailed descriptions of the study participants and by the use of prompts during the interviews to probe and clarify their statements and solutions.

Concerns about generalizability in qualitative research are often expressed in terms of transferability. A number of researchers have attempted to establish criteria and develop conceptual frameworks that address the generalizability of qualitative results. For example, Auerbach & Silverman (2003) distinguish between two kinds of results in qualitative research—abstract theory and specific theory. Specific theory, or specific patterns and themes that are identified in the data, might be transferable to similar

populations and settings or might not be transferable at all. However, more abstract or "holistic" results have the potential to be applied broadly to other situations. The potential transferability of findings from this study was discussed earlier in this chapter. While generalization to other specific populations of pre-service teachers may not be realistic, it is hoped that results from this research will be broadly useful in preparing secondary teachers to encourage students' construction and critique of viable arguments.

The validity of qualitative research can be improved by *triangulating* multiple data sources or multiple methods of analysis. In this study, triangulation is not a feasible strategy for data collection, as the task-based interviews provide the sole source of data. However, several measures have been taken to assure validity of the interview protocol and tasks. First, most of the tasks used with students have been effectively used in previous studies of proof and reasoning. Second, the items were tested in a pilot study with similar participants, and the lessons learned have resulted in improvements to the interview protocol. Third, the two evaluation tasks and the researcher-created "student responses" that go with them were reviewed by a mathematician and a mathematics educator to confirm the researcher's perception of their position on the five-level framework.

It was originally suggested that the researcher's faculty advisor would apply the data analysis rubrics to selected data to validate the researcher's coding process. However, as data analysis proceeded it was determined that the external coding system represented by the rubrics was adequately trustworthy. First, the rubrics were based on well-established research findings. Second, the researcher possessed extensive knowledge

of the literature on proof and reasoning. Instead, the researchers' faculty advisor focused on validating the internal coding system. She read and partially coded one of the interview transcripts and provided expertise on qualitative coding procedures. Frequent discussions served as a basis for validating the coding and analysis process without a detailed side-by-side comparison of themes and individual codes.

The method of triangulation can also be applied to the process of data analysis. Data from this study were analyzed from a number of perspectives: both vertically by case and horizontally across cases; by comparing participants' written work to their verbal statements; and by looking for inconsistencies among how the participants interpret, construct, and evaluate arguments.

Conclusion

This study provides an opportunity to investigate how the mathematical preparation of secondary teachers intersects with pre-service teachers' understanding of argumentation in the context of the newly adopted Common Core State Standards for Mathematical Practice. It combines a well-established research methodology, the clinical or task-based interview, with a highly regarded learning strategy, examination of student work. Finally, it expands on the current knowledge base about students' perceptions of proof and reasoning, and has the potential to inform teacher preparation in the new era of the Common Core.

CHAPTER 4

RESULTS

This research study explores how pre-service secondary mathematics teachers understand and interpret the term "viable argument," how they attempt to construct viable arguments, and how they evaluate the viability of mathematical arguments constructed by others. These practices were investigated through a series of task-based interviews with five pre-service teachers nearing completion of a teacher preparation program. This chapter reports findings based on those interviews and subsequent data analysis.

By way of introduction, a restatement of the research questions is followed by a rationale for how the data was collected and reported to best answer the research questions. A description is then presented for each participant in the study to provide an in-depth individual portrait of each case. These portraits include descriptions of the participants' background, their perceived mathematical proficiency, and their demeanor during the interviews including any attitudes or strategies they consistently demonstrated or employed.

In the following sections, individual results are reported for each of the three major categories of understanding, constructing, and evaluating viable arguments. Each section concludes with a cross-case analysis discussing similarities and differences across all five subjects. Finally, broad findings are reported on how this sample of pre-service teachers interfaces with the notion of viable argument through defining, constructing, and evaluating viable arguments.

Purpose of the Study

This research study set out to examine pre-service secondary teachers' understanding, interpretation, and application of "viable argument" as it is introduced in the Common Core State Standards. Extensive task-based interviews were conducted with the five junior/senior students in a secondary mathematics teacher preparation program in a midsize public university in the Rocky Mountain region. All five students were enrolled in the last of three mathematics courses that address proof, reasoning, and mathematical argumentation. The research questions are:

1. How do pre-service teachers define the notion of "viable argument"? Is their notion consistent with "viable argument" as defined by the Common Core Standards?
2. What strategies do pre-service teachers employ when asked to construct viable arguments?

   2a: Do they build a logical progression of statements to explore the truth of conjectures?

   2b: Do they use examples constructively in their arguments?
3. What characteristics do pre-service teachers assess to critique the reasoning of others?

   3a: Do they distinguish between correct logic and logic that is flawed?

   3b: Do they look for ways to clarify or improve the arguments of others?

   3c: Do they recognize and value constructive uses of examples?

Rationale for Data Collection. Analysis, and Reporting

"Viable argument" was first introduced as mathematical terminology with the Common Core State Standards in 2011, which makes it a relatively new term with no well-understood meanings or connotations. From the description given for Mathematical Practice 3 in the Common Core State Standards, it seems plausible that viable argument is attempting to describe a mathematical framework that emphasizes justifying results in a manner similar to proof, but with less emphasis on the rigor and formalism typically associated with the process of proving statements in mathematics. However, the distinction between a proof and a viable argument is not easy to pinpoint, as the CCSS describes viable arguments in terms that are often associated with the process of proving such as using "stated assumptions, definitions, and previously established results" and building a "logical progression of statements." The researcher stipulates that the notion of viable argument has no clear definition, and that the line separating it from proof is murky. The meaning of the term is further clouded by how both the pre-service teachers and the researcher interpret the phase. These facts had to be considered in designing how data was to be collected, analyzed, and reported as described below.

It was essential that the participants in the study had multiple opportunities to express their understandings about what it means to create viable arguments and to identify features of arguments that would make them viable. For this reason, the choice was made to collect data for this study through a carefully sequenced task-based interview. The interview began by exposing a subject's foundational knowledge of proof and argumentation. It then engaged each subject in discourse, problem solving, and

analysis of student work to provide multiple lenses from which to view the subject's understanding of viable argument. In this chapter, results are discussed in terms of each lens but also synthesized to accurately portray each subject.

It was equally important for the researcher to suspend his own views in order to build an unbiased portrait of each participant's thinking. This was partly accomplished by closely examining the researcher's own interpretations of proof and viable argument in Chapter 3. Bias was further reduced by the interview protocol, which allowed the participants to express their views in multiple contexts. Finally, the combination and repetition of conducting face-to-face interviews, examining video recordings, and coding written interview transcripts allowed the researcher to accurately deconstruct participants' thinking and to examine this thinking both within and across cases.

In general, findings reported in this chapter do not arise from single statements by the participants, but from a synthesis of their words and actions at various stages of the interview. During the coding process, the researcher began by identifying and grouping distinct words and phrases from each interview. Eventually, codes from several different stages of the interview were combined and collapsed to characterize how a given subject understands, constructs, and evaluates viable arguments. Wherever possible, quotes are used to support the researcher's interpretation; however, a finding that is derived from several statements during the interview is not always easy to represent in a direct quote.

Finally, examining how learners think about proof and argumentation includes making connections to the body of knowledge and research in that area. This study is not intended to create new theory. Instead, it is an effort to connect the recent focus on viable

argument in school mathematics, and pre-service teachers' understanding of that term, to the existing knowledge base about proof, reasoning, and argumentation. In order to make those connections, a report of findings must also include the researcher's interpretations in order to situate the participants' responses in the context of the literature on proofs and arguments. In that sense, this chapter on "results" includes both findings and interpretation. Conclusions are reserved for the final chapter.

Review of Participants and Program

The first steps of the data collection process began with developing a well-defined picture of the mathematical preparation of the participants related to argumentation of proof. The following is a brief synopsis of the teacher preparation program and an outline of the participant demographics.

The teacher education program that forms the backdrop for this study and the process of selecting participants were described in detail in Chapter 3. However, it is worth restating that the program is aligned with the recommendations of the College Board of Mathematical Sciences and the standards and practices outlined in the Common Core State Standards. The participants in the study had sufficient exposure to proof, reasoning, and argumentation to suppose a reasonable understanding of proof procedures and concepts such as proof by induction. Two instructors of the proof-related courses were asked to comment on the mathematical proficiency of the eight subjects based on their performance in prior or current coursework. The instructors individually determined a ranking as high achieving, average achieving, or low achieving, then compared and discussed their rankings until a consensus was reached. The final rankings included

students at all three achievement levels. The rankings of the five students who actually participated in the study reflected those of the larger sample, although more participants ranked towards the lower end of the rankings would have been desirable and an intra-group comparison was not possible with the obtained sample.

Individual Profiles

In this section each case is examined in more detail, resulting in a series of profiles of the participants. The profile for each case will be similarly structured, while still allowing for unique characteristics and informative differences to be highlighted. For each case, the profile will describe the participant's educational background and how the faculty ranked his or her mathematical proficiency. General demeanor during the interview, the researchers' perception of the participants' mathematical proficiency and confidence level, and their self-reported familiarity with the Common Core State Standards will be addressed. These profiles introduce the five participants whose stories populate this chapter: David, Bob, Bridget, Sam, and Linda. The profiles also provide a context for how each subject responded during the interview and engaged in the different tasks presented in the interview.

David

David was a junior in the secondary mathematics teaching program and had completed most of his mathematics coursework. In addition to the courses of interest in this research, he had also completed Linear Algebra and Introduction to Analysis. He was characterized as an average-achieving student by the faculty members who were asked to

rate the students prior to the interviews. Prior to completing the Calculus sequence, David took several lower level classes including College Algebra and Pre-Calculus. Based on his responses to the two construction items, the researcher chose to not ask David the additional construction item from Calculus.

David stressed the importance of proof, reasoning, and argumentation in school mathematics, and it was evident that he considered this to be a cornerstone of mathematics education. He expressed familiarity with the Common Core State Standards and stated an interest in the research project and the different tasks. He particularly mentioned the opportunity the interview provided to think about terminology he had not thought about before. At the conclusion of the interview he explicitly stated that he found it interesting and asked about the availability of the results once the research is complete.

David solved one of the construction tasks readily enough, but struggled with constructing an argument for both the angle sum problem and the transitivity of division problem. He was ultimately not able to construct an argument to his own satisfaction for either of these problems. He expressed frustration over this, stating that he felt he should be able to solve the problems. He mentioned that he did not do very well in Introduction to Analysis and that he had regrets about this. Based on these statements it seemed that David was not very mathematically confident, and his lack of confidence seemed to affect his performance on the construction tasks.

Overall, David's demeanor in the interview supported the notion that he was not very confident in his proficiencies with argument construction and that this lack of confidence affected his performance. He expressed a lot of hesitation, with frequent long

pauses as he was trying to solve problems. Statements like “I am not sure,” “I don’t see how I would start,” and “which probably wouldn’t help me” testified to his lack of confidence in both his own ability to construct a viable argument and his chosen approaches. When doing the angle sum problem, he did recognize a method that could produce an argument: “I was hoping that I can make the connection between the three angles to tell me something about this top angle,” but he was unable to find the connection and furthermore seemed to lack the confidence that he could find it: “I still feel that I would hit a wall at some point—like I need something else.”

The researcher’s impression was that David’s inability to construct arguments for the two items was caused by nervousness and the context of being interviewed and recorded on video. On both tasks, David had ideas that would ultimately have produced viable arguments—finding a connection between the angles for the angle sum problem and using common factors for the divisibility problem—and it seems likely that he would have been able to develop these ideas under a different situation.

Bob

Bob was a junior in the general science broad-field teaching major with a minor in mathematics teaching. The mathematics requirements for his degree were therefore slightly less than those of the other participants, but he had completed all the coursework necessary to participate in the study. Bob was characterized as high achieving by both faculty members who assessed the participants’ proficiency levels. Bob did not take any lower level classes prior to entering the calculus sequence. Due to scheduling issues, Bob

was the only participant who did the two interview sessions consecutively without a break between them. He was asked to complete the harder calculus construction item.

Bob approached each task confidently and he appeared secure in his mathematical understanding and ability. He solved both of the construction tasks relatively easily and was therefore asked to complete the harder calculus construction item as well. He solved the initial angle sum problem confidently and was confident in his solution even though it did contain a minor flaw. Bob expressed some hesitancy with calculus, which he clearly thought was a hard subject, but he still readily accepted the challenge of a harder construction item.

Throughout the interview, Bob seemed to approach the tasks from a teaching context, several times referring to an audience as "they" and highlighting how he would address something in class and what prerequisite knowledge "they" would need. He was the only one of the participants that consistently approached the tasks in this fashion.

As a broad-field science major, Bob's science background was very evident in the interviews as he talked about the power of observations in forming conjectures, expressing views more typical of scientific than mathematical method. It is evident that his knowledge of scientific method influenced his understanding of mathematical proof and viable arguments.

When asked about his familiarity with the Common Core State Standards, Bob stated that he was "very familiar with how they work and how to apply them, I don't really know what they are." Presumably Bob had used the CCSS frequently in classes and was familiar with how to compare mathematics problems to particular standards at

particular grade levels. This would explain how he knew "how to apply" the standards. However, it is possible that Bob did not know that the CCSS are an attempt at a national curriculum currently adopted by 45 states, how they came about or who wrote them, and other political and mathematical background information.

Bridget

Bridget was a secondary mathematics teaching major with senior standing who was characterized as high achieving by one of the faculty members and high to mid achieving by the other. In addition to the three classes of interest in this study she had also completed Introduction to Linear Algebra, Introduction to Analysis I, and Geometry, Measurement, and Data in the Middle Grades. Based on this coursework, Bridget had a solid background in mathematical proof, reasoning, and argumentation as well as documents and teacher recommendations supporting the CCSS. Bridget completed the harder construction item from calculus.

The interview sessions with Bridget took place on two consecutive days. In the first interview session, she seemed a little uncomfortable with the interview setting and the video recording, but this uncertainty diminished as the interview progressed. Bridget was interested in the purpose behind the questions, on a couple of occasions asking the researcher if her answers were satisfactory. When the researcher re-emphasized that the purpose of the interview was to uncover the participants' thoughts and that there were no right or wrong answers to the questions, she accepted this explanation and seemed more confident in her responses.

Bridget showed evidence of having a very strong mathematical understanding relating to proof, reasoning, and argumentation, and she was not much challenged by the two construction items. She also came very close to providing a proof for the more challenging calculus item, but she missed a connecting step and she decided to move on to attempt a different solution strategy. Despite her failure to produce a proof, Bridget solution to the calculus item was the closest any of the participants came to proving this item without any input from the researcher and with more time it is very likely that she would have made the connection required to prove the statement. Overall, Bridget's behavior gave the impression that her mathematical understanding was strong, but her mathematical confidence was not.

Bridget also displayed a mathematical curiosity as she was the only one of the participants that asked the researcher how to prove one of the construction item, when she was dissatisfied with her own solution attempts. She also expressed interest in the research study and how the study would inform the teacher preparation program. She stated familiarity with the Common Core State Standards, saying that she was most familiar with the Mathematical Practices.

Sam

Sam was a junior in the secondary mathematics teaching program. He was rated as an average-achieving student by one faculty member and as average to high-achieving by the other. Prior to entering the calculus sequence Sam had completed several prerequisite courses including College Algebra and Pre-Calculus. He had also completed Introduction to Linear Algebra. Due to scheduling issues and the interruption of spring

break, close to two weeks separated the two interview sessions with Sam. It is possible that this impacted the interview as a whole even though the data does not reflect any immediate impacts. Due to the relative ease with which Sam answered the two construction items, he was asked the additional more challenging calculus item.

Sam was "roughly familiar" with the Common Core State Standards. In the interview he was very expressive, using many hand movements for added emphasis as well as a lot of non-verbal communication. He appeared confident and he answered questions readily with little to no hesitation. Sam's confidence was most apparent when he remembered a solution process. He immediately provided an argument for the angle sum problem stating that he remembered how to do it. When asked if there were other methods, he recalled seeing another approach but was not able to provide a solution using this method, and he quickly gave up, returning to provide more detail on his initial response.

Sam's behavior during the interview implied that his confidence bordered on over-confidence. This may have influenced his decision-making and his inability to change his initial assessment of a mathematical situation. For example, Sam's assessment of one of the student works items differed from the assessment of both the researcher and the mathematics faculty member asked to evaluate the item. Even though the researcher offered very specific and leading questions as to the merit of this particular item, Sam was unwilling to change his initial assessment.

Linda

Linda was a senior with a double major in business and secondary mathematics teaching. As a senior she had completed Mathematical Modeling for Teachers as well as History of Mathematics. As part of her business degree, she had also completed a Survey of Calculus course. Linda was rated as a low achieving student by one of the faculty members, and a low to average-achieving student by the other.

Linda seemed to not be as mathematically proficient as the other participants, and this was reflected in her demeanor and her confidence during the interview. She seemed hesitant and insecure in her statements and she was not confident in completing the construction and evaluation tasks, going back and forth on whether or not arguments were viable. She struggled with the fact that the second construction item is the transitive law applied to divisibility. Linda wanted to say that this was true by the definition of transitivity, not recognizing that it is a property that needs to be proven. She was not able to provide a solid justification for this second construction item and as a result the researcher decided not to give her the additional more challenging construction item taken from calculus.

In the remainder of this chapter, data is reported for the three major phases of the task-based interview: defining viable arguments, constructing viable arguments, and critiquing the viability of others' arguments. Results are reported separately for each phase. In each section, individual results for the five subjects are reported first, followed by a summary of cross-case analysis of all five subjects. This discussion begins with an

attempt to further introduce the five subjects by reporting their responses to an initial warm-up task.

Analysis of the Initial Task

In the first task of the interview, participants were asked to provide an argument supporting the true statement that the sum of the angles in a triangle is 180 degrees. The purpose of this task was to see what approach the pre-service teachers would take when asked to support an argument. The researcher anticipated that many of the participants would attempt to prove the statement formally. This would provide an opportunity to stress early in the interview that the focus was on viable arguments as introduced by the CCSS and not on formal proofs, which had been the focus of many of the participants' recent mathematics classes. To further open the task to informal solutions, the participants were told that they should treat the task as if they had access to whatever resources they might find useful, such as manipulatives, tools, and software. The responses to this initial task are summarized below for each case, before general trends across the cases are highlighted.

David

David expressed that he found the task difficult, because it was such a basic fact in mathematics. He was more used to using it as a prior result in other problems rather than providing a justification for it, stating that "It is…challenging to me because I think of it as such a baseline…I would assume this to be true and then build on it." He tried several avenues to prove the statement, but failed to find one that met his criteria. All of

his attempts were based on trying to prove the conjecture, and he initially offered no alternative solution that would explain the result or even demonstrate its truth. He explicitly stated that he was trying to prove the statement.

When asked about other methods he turned to things familiar and "things I understand" because obviously "I am not going to be able to provide a good proof if I don't know what I am talking about." He either attempted or considered using "right angle" trigonometry because he liked it and using circumcircles because they had done a lot of work with those in his Modern Geometry class. He discarded both of these when he failed to see an avenue to pursue.

It is worth noting that one of the approaches David mentioned would have provided a proof had he been able to complete it. At one point he drew a line parallel to the base of the triangle through the third vertex (Figure 7). However, he failed to find the idea of alternate interior angles that would provide him with the congruencies necessary to complete the proof. He also stated that he did not think he would be able to find a way to "connect the angles."

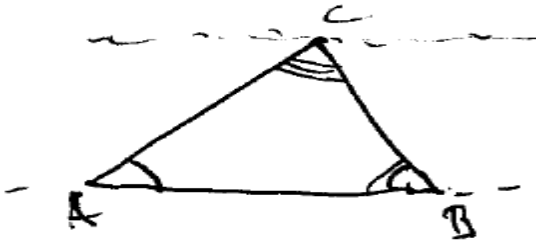


Figure 7: David's Attempted Solution to the Angle Sum Problem

When asked how he would convince a 6th grade class of the result, David mentioned using dynamic software to create a triangle and display the angle sum as the vertices are moved around. However, he did not necessarily believe everybody would be convinced by an approach using dynamic software. He was not sure if the absence of a counterexample would be entirely convincing, but stated that he himself would be convinced of it. He also mentioned that it would be possible to work backwards from a rectangle (i.e., drawing a diagonal creating two triangles), but was unsure about whether or not the angle sum of a rectangle could be used as a prior result.

From his responses it is evident that David interpreted the question, which was phrased "provide an argument in support of the following statement," as "prove the statement." He immediately attempted to provide an argument that would have classified as a mathematical proof had he completed it, and he continued his attempts as proofs when asked for alternate ways of solving the problem. Only when asked how he would convince somebody in a lower grade classroom did he talk about other approaches less formal than proof.

Bob

Like David, Bob also tried to prove the statement about the sum of angles in a triangle, and he provided a general argument that would be rated as proof had it been completely correct. Bob's solution was based on creating a parallelogram by using two sets of parallel lines and then drawing one of the diagonals, but his argument contained a flawed assumption that the diagonal would be an angle bisector (see Figure 11). This

flaw is easily corrected using alternate interior angles, and the researcher felt certain Bob would be able to fix it had the flaw been pointed out to him.

When asked to provide an alternate solution to the angle sum problem, Bob attempted multiple solution strategies based on his recollection of what had been covered in classes. None of these attempts was successful, and all of them involved trying to prove the statement. During these attempts his language changed noticeably. In his original solution, albeit not 100% correct, he used phrases such as "I know this" and "Now we know that" and he did not hesitate in providing his solution. When seeking alternate solutions, Bob paused more frequently and his statements were less confident: "Where do I go with that?...I don't really remember where.".

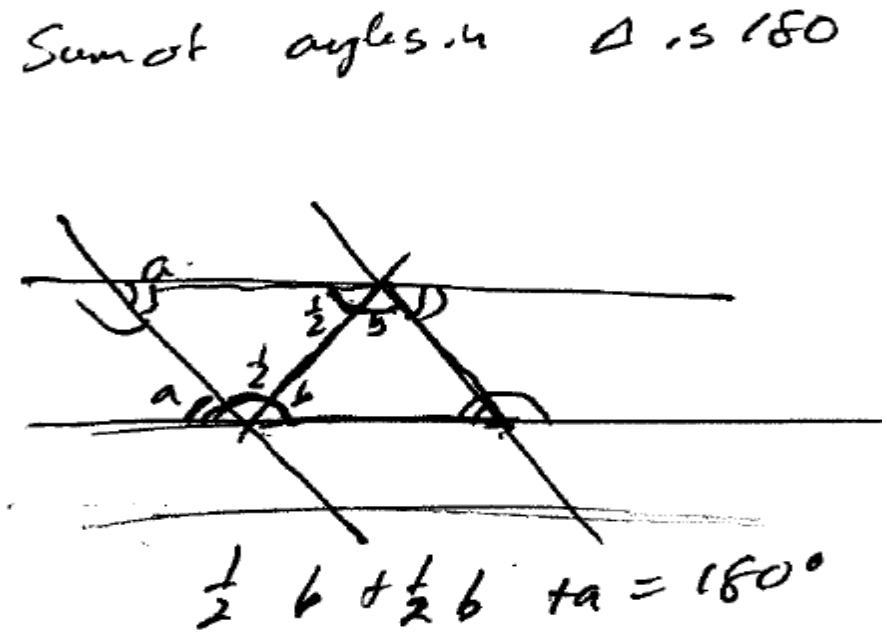


Figure 8: Bob's Written Part of the Solution to the Angle Sum Problem

Bob did not mention using dynamic software or manipulatives in his different attempts at the angle sum task. When asked about convincing rather than proving the statement, Bob answered that he would convince somebody by considering a square and drawing the diagonal to create two right triangles and that he valued the use of concrete examples, especially in a teaching context. He also described a heuristic argument starting with an equilateral triangle and explained how if you increase one angle, then one of the other angles will have to decrease and the sum will remain constant. It is interesting that Bob was currently enrolled in a class that emphasized the use of dynamic software, but even though he mentioned changing the angle measures in triangles he did not mention the use of dynamic software.

Bridget

Bridget was the only participant who did not immediately attempt to prove that the angle sum is 180 degrees. Instead, she opted to describe how to use dynamic software to show that the statement is true, specifically dragging a vertex in a triangle down to the base to create a line. Bridget stated that this would be an intuitive approach, but seemed hesitant and not entirely happy with her method. When she returned to this approach later in the interview, she commented that it was not very convincing, but did not seem to be interested in trying to make it more convincing. In a way her response was reminiscent of the professors in Raman's study (2003), who were content with describing a key idea knowing that they could formalize the argument if needed. Bridget knew that she could use dynamic software to show that the statement is true, even if the exact method escaped her in the interview.

Bridget's response is unique in that she did not immediately attempt a proof of the statement, but instead thought of a way to show that the statement was true. This indicates that to her, supporting the truth of a statement is not necessarily equivalent to proving the statement. Furthermore, Bridget was well aware of how to prove the statement; immediately after providing her dynamic software solution, Bridget asked if the researcher wanted another solution to the statement and proceeded to easily prove the statement using the standard method of constructing a line parallel to the base through the third vertex and showing that all three angles together form a line by considering alternate interior angles.

Also of interest is that after completing a verbal proof with the aid of an illustration that the researcher readily accepted as a proof, Bridget asked if the researcher wanted her to "write out a proof of it." This suggests she did not consider a verbal argument to be a proof and that it needed to be formally written out in order to qualify as such. She referred to her argument as a "pretty good argument" but not as a proof, most likely because it lacked the formalism and rigor she associated with a mathematical proof.

Overall it appears that Bridget had a fairly sophisticated understanding of proof and had no trouble in producing a proof. Yet she did not necessarily consider a proof to be the best argument if the goal is simply to support the truth of a statement, whereas all the other participants immediately attempted to prove the statement. Furthermore, Bridget clearly regarded proofs to be written formal arguments and did not view a verbal argument explaining why a statement is true as proof regardless of how good the

argument might be. In this way she was distinguishing between explaining or convincing on one hand, and proof on the other.

When prompted, Bridget attempted a couple other ways to show that the angle sum statement was true, particularly by drawing a rectangle and its diagonals, but she recognized that this method would not be general since it covers only right triangles. She suggested using dynamic software to convince a 5th grader and stated that she would use different methods to convince a 5th grader and a mathematician, indicating that what she regarded as a good or convincing argument depended on the audience.

Sam

For the angle sum task, Sam produced a proof using parallel lines and alternate interior angles, stating that he chose this approach because they had done it this way in class. Like Bob, Sam was confident in his result as he clearly understood and remembered the procedure.

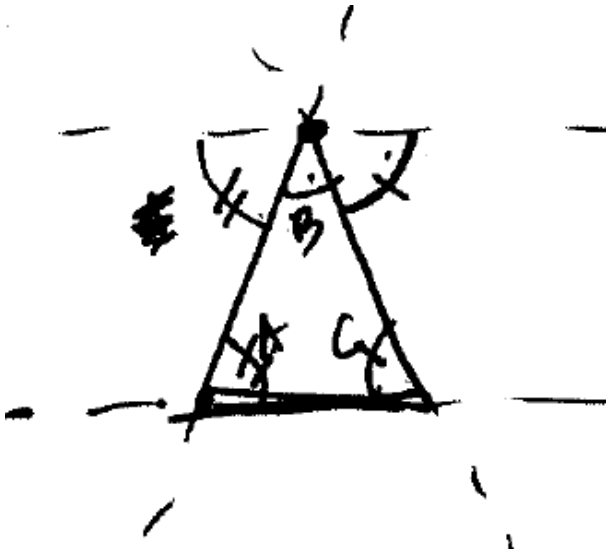


Figure 9: Sam's Written Part of the Solution to the Angle Sum Problem.

Sam mentioned another method of proving the statement that he remembered from high school, but he was not able to get to a solution using this method. In this respect Sam's response was similar to Bob's; the alternate solutions were not internalized in the same manner as the original solution he produced and he quickly gave up, returning to justifying the result he had internalized.

When asked about how to just "convince" somebody as opposed to proving the statement, Sam said he would use dynamic software and that this would be one way to "show it to somebody who doubts it." The implications are that Sam equated the task of producing an argument with the task of producing a proof, but when asked to convince somebody, he preferred using dynamic software to illustrate that the statement is in fact true, perhaps because of the mathematical sophistication required to understand proofs. In his view, conviction and proof were not the same and different contexts called for different measures.

Linda

Linda produced the outline of a proof using parallel lines and alternate interior angles, stating that they had been talking about congruent angles in Modern Geometry and that she was therefore comfortable with using this approach. Like Bob and Sam, she turned to knowledge that she had internalized. Linda's solution to the problem is shown in Figure 10.

When asked if there were any other methods she could use, Linda interpreted this as looking for another way to prove the statement. She suggested maybe using trigonometry, but stated that she "is awful" at trigonometry and was not confident she

could find a solution using this approach. Linda was still interpreting the task as asking her to prove the statement, and she continued to do so even when the researcher asked her if this was the method she would use with a friend or a 5th grader. Only when it was pointed out that she merely needed to produce an argument, not a proof, did she let go of the notion that she had to prove the statement. She asked "Just explaining it to a regular person?" implying that proving and explaining were not the same thing. Despite this clarification, she did not have a way to provide a good explanation to a 5th grader other than making her initial response less formal. Using software did not occur to her, especially surprising since Linda was the only participant who had completed the mathematical modeling class and was currently enrolled in Modern Geometry, the two classes in the teacher preparation program that emphasize the use of technology.

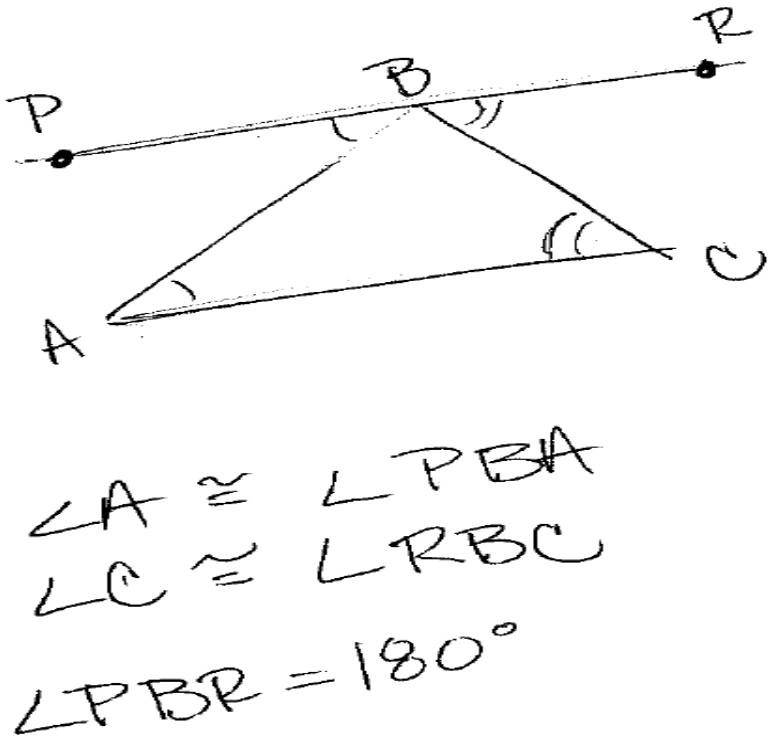


Figure 10: Linda's Written Part of the Solution to the Angle Sum Problem

Analysis across Cases

Data from the angle sum task make it evident that these pre-service teachers as a group interpreted the task of providing support for a true statement as a proof task. Of the five participants, four initially attempted to prove the statement, whereas the fifth asked if the researcher wanted a proof of the statement after providing a statement using dynamic software. Among the attempted proofs, three of the participants successfully produced a proof of the statement; the fourth participant provided a flawed argument that could be easily corrected to a proof and the fifth participant was unable to provide a satisfactory argument.

It is possible that "performance pressure" in recording their thinking in an interview contributed to the participants' focus on proving the angle sum statement. Still, it is interesting that the participants were not more inclined to use dynamic software in support of their arguments, particularly since it was a geometry task and the Modern Geometry class relies on using dynamic software. Even when the researcher emphasized that the question only asked for support of the argument and not necessarily a proof, the participants favored attempts at proofs, and most of the alternate solutions offered by the participants were attempted proofs rather than less formal arguments. This may indicate that the participants have ample training in mathematical proof, whereas they may not have much training with other types of arguments. It may also indicate that training in justifying mathematical results focuses more on proof production than creating less formal arguments. It is also surprising that none of the participants mentioned the method

of using scissors to cut out the angles of a triangle and line them up to form a straight line, since this method was mentioned by two of the three participants in the pilot study.

All five subjects relied on their prior knowledge of a particular approach when solving the problem. With the exception of David, who could not provide a satisfactory solution, all the participants remembered a solution and used this as a basis for their arguments. By asking if there was another way to solve the problem, the researcher's intention was to see if the students could produce arguments that were not proofs. However, the participants all interpreted the question as asking if they knew a second proof of the statement. Only when the researcher asked about convincing arguments did some of the participants mention other methods. This behavior emphasizes the participants' proclivity to interpret the question as a request to produce a proof as well as their reliance on memorization of a solution pattern in solving the problem rather than their own mathematical understanding.

Analysis of Viable Argument Tasks

The following section presents the participants' responses to questions about the nature of viable arguments. The emphasis is on how the participants interpret the term "viable argument" and what they consider to be the role of viable arguments in school mathematics. The results include a comparison to the description of viable argument found in the Common Core State Standards and a discussion of how the participants see viable arguments as different from or similar to proof. From the analysis of the participants' responses, ten different themes emerged based on the internal and external

coding schemes. Tables outlining these themes and some key phrases from the coded data that supported the establishment of the themes will be presented for each participant. The cases will first be treated individually in the same order as in the previous section, followed by a cross-case analysis where common themes will be identified.

David

For David, viable arguments display mathematical reasoning and communicate mathematical understanding to others. A viable argument should contain a "mathematical approach" where students think from "a mathematical perspective and try to use definitions and logic." David's description of a viable argument was very close to the description given in the CCSS; prior to being handed the CCSS statement, he mentioned that a viable argument needed to contain a logical progression "to get from point A to point B" and that it should be built on the use of definitions and logic. In David's opinion this was at the heart of mathematics, and he regarded attempts at producing an argument, explaining reasoning, and displaying understanding as more important than successfully producing a solution. He viewed viable arguments as adhering to certain guidelines, such as using definitions and logic. Following these guidelines was a goal in itself: "Whether or not they actually succeed wouldn't really be so important to me, but they are at least understanding the procedure and the route." For David, creating a viable argument was not just a mathematical practice, it also involved mathematical content that followed a procedure and a route.

David mentioned that viable arguments are closely related to proofs, but need not contain the formality and rigor associated with proof, referring to a "form of proof" in

which logical steps and a logical progression are important. He understood and commented on that relationship with proof: "I think a viable argument is less thorough than a proof but I can see how they are the same."

When asked about the difference between a proof and a viable argument, David noted that context is important. He stated that a proof was to convince a mathematician, whereas a viable argument was almost like convincing oneself. His words echo the oft-quoted description of proofs as containing a sequence where you first convince yourself, then convince a friend, and then convince an enemy (Mason, Burton, & Stacey, 1982). David also expressed that viable arguments are more appropriate for school mathematics than proof: "I do not expect my sixth graders to lay out a formal proof, but if they can give me a solid argument, a viable argument, then I would think that is…a really great step towards making a proof." For David, viable argument is a precursor to formal proof; they are similar in that they use logic and display mathematical thinking, but viable arguments are age-appropriate and do not have to adhere to the rigor of a mathematical proof.

According to David the foremost role of viable arguments in school mathematics was to display mathematical thinking and understanding. "I think a viable argument is strong proof of mathematical thinking." A viable argument showed "why it works." Additional roles were to learn how to communicate mathematically, so that someone who "isn't in your brain is understanding what you are saying," to provide explanation for results, and to serve as an introduction to logic and mathematical methods to make the transition to proof classes easier. David regarded viable arguments as important in that

they display students' understanding of why something is true as opposed to rote memorization.

An overview of themes that arose in the analysis of David's responses to questions concerning viable argument is given in Table 4. Themes that were not prevalent in David's responses are left blank.

Table 4: Coding of David's Responses to Definition Items

Viable Argument Characteristics and Features		
Participant	Themes	Quoted Key Words and Phrases
David	Comparison to Proof	form of proof; not go as far as proof, less thorough; almost the same…depending on context; step towards making it a proof
	Mathematical Approach	logical steps; use definitions; use sound reasoning; get from beginning to end; strong proof of mathematical thinking
	Explanation	this is why it works
	Understanding	they are…understanding the procedure and the route; make sense to them; they would understand it
	Convincing	convincing a mathematician versus …convincing yourself; try to convince me
	Generality	make arbitrary
	Correctness	whether or not they succeed would not be important
	Student-Centered	if they can give me an argument; they can…
	Context	depending on context

Bob

Initially Bob stated that a viable argument is based on "an observation with some sort of explanation." He mentioned that a viable argument would need to incorporate an explanation of a phenomenon. He repeated this assertion twice, first saying that a viable argument should have "some sort of even an elementary explanation" and later adding that the essence of a viable argument is to explain why something works. For Bob, an explanation does not have to be detailed or formal; in fact, one of the things that separates viable arguments from proof is that in proof theorems must be cited, whereas in viable arguments it is enough to explain that "this works because of this" without necessarily formally citing the theorem. Furthermore, an explanation might consist of multiple observations of a phenomenon: "[T]heir argument would be strong to me because they have observed it in many different trials." In his discussion, Bob specifically mentioned dynamic software as a method for constructing viable arguments, as dynamic software has the ability to generate a multitude of examples.

Although Bob explicitly and repeatedly stated that a viable argument can be based on multiple observations, he was clear that a single observation was not sufficient for a viable argument. He used the example of only considering an equilateral triangle as insufficient for creating a viable argument since not all triangles have 60-degree angles. Observations that serve as the basis for viable arguments had to be either recurring observations or an observation with some sort of a generic quality to it.

These statements suggest that Bob was aware of the limitations of empirical evidence in mathematics. On the other hand, his repeated assertions that multiple

observations, which essentially are empirical evidence, can constitute a viable argument seem to indicate that he was not aware of these limitations. This apparent contradiction is explained by two factors. First, for Bob, viable arguments were tied to their ability to convince. Bob explicitly stated that using parallelograms to explain the angle sum result to second graders would not be able to convince them and would therefore not be viable. On the other hand, multiple observations would constitute a "strong" argument because it would convince them of the result and would therefore constitute a viable argument.

Second, as a broad-field science major Bob was heavily influenced by scientific methodologies in his approach to mathematical proof and reasoning. In general, the scientific method consists of forming testable hypotheses. Although a hypothesis may never be formally proven, it can be falsified, and recurring observations and testing that fail to falsify a hypothesis are seen as evidence in support of the hypothesis. In science, using dynamic software to test the hypothesis that the sum of the angles in a triangle is 180 degrees is sound methodology. In mathematics, the same approach, although convincing, does nothing to prove the statement true. Bob specifically mentioned that he thinks of viable arguments as hypotheses, and proofs as the explanations for why these statements are true. He understood that empirical evidence did not mathematically prove a conjecture, but empirical evidence was sufficient to create conviction as long as the examples had some generic quality or are repeated.

The difference between proof and viable argument was to Bob a difference in the formality, rigor, and detail with which the result is explained. He stated more than once that a proof requires citing theorems, and he referred to his initial argument for the angle

sum as a "borderline proof" that could be a proof if he had cited the theorems he used. Bob stated that "A viable argument…is based upon just observation….I have observed this and I think this is why, whereas a proof goes into 'this is exactly why it works.'" It is clear that Bob regarded proof as a formal write-up supporting the truth of a statement.

Bob interpreted the CCSS description of viable arguments as a basis for inquiry and active learning. He stated that observations can be used to form conjectures and students can then construct arguments based on previous results in a student-centered "active learning" environment. Bob valued argument construction as a means to make instruction more student-centered, and he contrasted this student-centered approach with a lecture-based learning situation where teachers are "going back and teaching them theorems." An interesting exchange occurred when Bob was asked about the role of viable arguments in school mathematics. He replied that teachers do not have time to teach all the proofs of theorems that they encounter, and teaching viable arguments will free up time to focus on other aspects of mathematics. He saw viable arguments as a vehicle to construct new knowledge from previous knowledge; "I know multiplication and I know angle addition, so I can apply those things and construct an argument for this." He added, "I think there is not enough time in the year to go back just to introduce them to things just to prove." Inherent in this seems to be the view that proof is a teacher-game (Reid & Knipping, 2010) where students prove statements they already know to be true. There was also an implicit view that viable arguments promote understanding whereas proof is an exercise in formally writing up results apparent in Bob's response. This is evident in the following quote:

> But if it is a viable argument that has the correct answer, that is a good strong thing for them to just like grasp on to and bring back when they need it in later math classes or in college or anything like that….They have that good quick one class period understanding where they constructed their own argument for this, and they remember that because they did it themselves. They'll have that rather than going over a proof for three days in a row step by step why this works….It is going to take a lot more time with basically the same effect in the end, in that they understand [the result].

Bob's comment reflects the notion that proof is a time-consuming formal write-up of results that can more easily be understood through explanations. The preference for viable arguments is thus a pragmatic one: viable arguments produce better explanations faster. He sees the construction of a viable argument as student-centered, whereas proof is centered on a teacher lecturing. Viable arguments are thus preferred as they embrace active learning practices. In the end both achieve understanding of the result, but because of its active learning aspects, viable arguments have higher potential for student retention of the material. Bob's description of viable arguments versus proofs is reminiscent of Tall and Vinner's (1981) distinction between concept image and concept definition, which in their words is "a distinction between the mathematical concepts as formally defined and the cognitive processes by which they are conceived" (Carpenter, et al 2004, p. 99).

A summary of the themes identified in Bob's responses are given in Table 5. Themes that did not appear in Bob's responses are left blank.

Table 5: Coding of Bob's Responses to Definition Items

Viable Argument Characteristics and Features		
Participant	Themes	Quoted Key Words and Phrases
Bob	Comparison to Proof	proof…flawless…cite this theorem; strong educated hypothesis versus exactly why it works; maybe they are missing a math theorem; borderline proof
	Mathematical Approach	
	Explanation	observation with explanation; have to have…elementary explanation; explained this works
	Understanding	a good concrete thing to hold on to
	Convince	I would be convinced; convince somebody of this
	Generality	many different trials; generic observation to every type; arguing various cases; recurring observations
	Correctness	
	Student-Centered	they construct; using what they already have; a student-led conversation; like an inquiry
	Context	

Bridget

Bridget expressed that she was not quite sure what was meant by the term "viable argument" and that she would have to consult others to fully understand what is meant by it. When provided with the lexicon definition of the term "viable," she interpreted viable argument to mean an explanation of the result. She was the only participant who

mentioned that a viable argument could be an argument that uses a key idea to explain why something is true, without necessarily writing it out. For Bridget, the explanation is the important part in a viable argument and not the format or "layout," which is important in proofs.

Bridget's initial differentiation of the terms proof and viable argument was partly based on the formality of the argument ("If they had wanted proof, they would have said proof, not viable argument") and partly based on the difficulty level associated with teaching proof and the necessary background information required for students. In this sense, she distinguished between learning about proof in general, which she thought was a good idea, and the act of actually writing a proof, which she did not think should be required of students.

When asked about characteristics of a viable argument, Bridget focused on its ability to explain. She felt it important that the steps in an argument are logical, cover all cases, and use prior results, but she regarded it as primarily an explanation. She believed that in school mathematics viable arguments promote and show student understanding of mathematical ideas; students could use viable arguments to display understanding of why something works as opposed to "just doing things." Bridget considered an argument's ability to explain why something works as vital to its viability. She regarded her initial argument using dynamic software to show that the angle sum of a triangle is 180 degrees as not viable because it lacked explanatory power. She thought that this argument was intuitive and that showed that "it works," but since it failed to explain why it works, this disqualified it as a viable argument.

Table 6: Coding of Bridget's Responses to Definition Items.

Viable Argument Characteristics and Features		
Participant	Themes	Quoted Key Words and Phrases
Bridget	Comparison to Proof	not a step by step proof; explain why it works…different to an actual proof; can produce it into something more laid out; less formal
	Mathematical Approach	follows logically; uses prior knowledge; key idea; looking for the key thing; why it works
	Explanation	explain why it works; just an explanation; more of an explanation; because of this, this happens
	Understanding	so they understand why it works
	Convincing	is convincing
	Generality	have considered most of the cases; can't just explain one case
	Correctness	
	Student-Centered	have them explain
	Context	depends on the content

Initially Bridget thought of viable arguments as less formal than proof, but after reading the CCSS description of viable arguments, she stated that proofs and viable arguments were essentially the same in high school mathematics and that high school students were expected to "prove things." She based this on the language used in the description with counterexamples and logical progressions. She saw this as an

implication for teaching because teachers would have to teach how to logically set up an argument using definitions and prior results. For Bridget, the difference between viable arguments in high-school settings and in lower grades was one of formality and rigor. After reading the CCSS she interpreted viable argument to mean proof in a high school setting and an explanation in lower grade settings. She saw the difference as a matter of content "because in high school you want them to use… what they already learned in middle school to… explain why things work."

Sam

To Sam the important characteristic of a viable argument was that it is explanatory. In order to construct a viable argument, a student needed to "explain their reasoning to some extent" and explain "where their process is coming from." He saw viable argument ranging from proof, which is very detailed and step-by-step, to short justification for why a process is applied and why this process is appropriate. What qualified as viable was also context dependent—he would expect college-level viable arguments to be more like proofs, whereas viable arguments did not have to be as formal and structured in school mathematics.

Sam saw the difference between viable arguments and proofs as one of formality. He stated that a proof follows from using definitions, axioms, and prior results, while a viable argument was an explanation that is not necessarily formal. In school mathematics, Sam regarded the purpose of viable argument as providing a justification for why a particular approach was used, as opposed to just "getting to an answer and doing more of a computational thing." He specifically mentioned using finger -counting as a visual

argument for why addition works as something he would consider to be viable in elementary grades. He further stated that he regarded the logical structure of mathematics as important at all levels, whether in college or in grade school, and that structure might take on more of a visual aspect in lower grades as long as it provided a justification for the result at hand.

Sam considered both of his angle sum arguments to be viable arguments. He considered the interior angle solution to be more of a rigorous proof, at least if it was done with a "more detailed write-up," whereas the dynamic software approach was viable because it allowed the argument to be made as general as possible with a visual representation. He recognized that the dynamic software was not a proof and could not exhaust the possible number of triangles, but he still considered it to be a viable argument based on its ability to convince someone that the statement is true. Unlike Bridget, Sam did not elaborate on how his dynamic software solution failed to explain why the angle sum in a triangle is 180 degrees, considering its ability to show a multitude of examples sufficient to qualify it as viable.

A summary of the themes that surfaced based on the analysis of Sam's responses to the viable argument items are reported in Table 7. Themes that were not supported by data are left blank.

Table 7: Coding of Sam's Responses to Definition Items

Viable Argument Characteristics and Features		
Participant	Themes	Quoted Key Words and Phrases
Sam	Comparison to Proof	definitely not a proof; not have to be quite as structured as proof; not in a formal sense; definitely not a proof
	Mathematical Approach	definitions…theorems…what you can do with them
	Explanation	be able to explain their reasoning; explaining where their process is coming from; short justification; brief explanation; show where they are coming from
	Understanding	you are understanding the material
	Convincing	
	Generality	they can apply what they know…to an infinite set of numbers; as generally as you can
	Correctness	
	Student-Centered	they can…
	Context	in college more proof; depends on audience finger -counting in younger ages

Linda

Linda regarded a viable argument as an argument that provides "the reasoning behind why they answered the problem in the way that they did." Viable arguments still use theorems that are known to be true, but the arguments themselves might not be formally written out with the rigor associated with proof. She noted a similarity with

proof in that it is important to be able to explain and justify the steps taken, but felt the term "proof" has connotations of rigor that the term "viable argument" does not. Nevertheless, she believed it important that a viable argument conveys the author's thought in a manner that makes it understandable to others. For Linda, a viable argument is able to both highlight the author's understanding and explain this in a manner that is understandable to others.

Linda stated that viable arguments are important to escape the "plug and chug" conception many students have of mathematics, and that argumentation permeate all of school mathematics. She thought viable arguments had the ability to promote understanding of procedures by showing why formulas or procedures are applied in order to solve problems. In her opinion, the difference between viable arguments in elementary school and high school is a difference in sophistication and subject matter, but the important part is that students get to explain why something works the way that it works.

When asked if the angle sum argument she constructed was viable, Linda seemed confused about the difference between a viable argument and a proof. She initially stated that her argument (Figure 10) was not viable because a viable argument depends on context— it is necessary to know knowledge base or what is known to be true in the classroom community. She believed it would become viable if she formalized it by explicitly stating the congruence properties used in her solution. She did not know how to answer when asked if this would technically make her solution a proof. She did change her solution to be considered viable when asked if it was viable in the setting of her geometry class or in the context of the interview. It seems clear that Linda struggled with

knowing what can be assumed as prior knowledge, and that prior knowledge influenced what she thought of as a viable argument.

An overview of the themes that emerged from Linda's discussion of viable argument appears in Table 8. Themes that did not emerge are left blank.

Table 8: Coding of Linda's Responses to Definition Items.

Viable Argument Characteristics and Features		
Participant	Themes	Quoted Key Words and Phrases
Linda	Comparison to Proof	not necessarily formally written out; not as formal
	Mathematical Approach	valid statements; prior knowledge; accuracy; theorems
	Explanation	show the reasoning behind; not just giving an answer; being able to back it up; being able to explain
	Understanding	show a deeper understanding of the steps you took; showing your understanding; translating understanding to others
	Convincing	
	Generality	
	Correctness	
	Student-Centered	they use…
	Context	more basic in elementary; level of difficulty

Analysis across Cases

Viable Arguments

All the participants mentioned the importance of explaining reasoning in mathematics, and saw a viable argument as an explanation of a result. The participants

used slightly different terms for this: Bridget, David, and Bob used the term explanation, Sam used justification, and Linda referred to the reasoning behind a result. Essentially they all saw a viable argument as an argument that outlines why a certain result holds or why a statement is true; however, what they regarded as an explanation somewhat differed. For David and Bridget, the argument needed to be logical and structured, whereas Bob and Sam thought that a viable argument could consist only of observations as long as they had a generic quality to them. The difference in interpretation was perhaps most evident in the fact that Sam accepted using dynamic software as sufficient to qualify as a viable argument, whereas Bridget disqualified the same approach because it only shows that the result is true but does not explain why.

In general, the participants' definition of viable argument aligns relatively well with the description outlined in the Common Core State Standards. All five participants focused on viable arguments as a means to justify conclusions, explain mathematical thinking and communicate mathematical understanding. Several of the participants also commented on how viable arguments are not fixed, but may increase in sophistication as students progress through the grades. To illustrate this, Sam regarded finger counting to add numbers as a viable argument. Both David and Bob seemed to express sentiments that a viable argument need not be correct. It is, however, not clear if by this they meant that logical flaws are admissible, if they were referring to oversights such as the fixable flaw in Bob's angle sum argument, or if they were talking about inaccuracies and errors in the form and presentation of the argument.

Relationship to Proof

All five participants compared viable arguments to proofs. There was a consensus that a viable argument is similar to a mathematical proof but lacks the formalism and the rigor associated with proof. It is evident that the participants viewed a mathematical argument as a formal, logically rigorous argument that has to adhere to certain forms of presentation. For instance, proofs were expected to be in a written format, as exemplified by Bridget's question if she should write out a proof after producing one verbally. In contrast, the participants' comments implied that a viable argument allows for more leeway both in content and in structure.

Role of Viable Argument in School Mathematics

As mentioned above, all five participants viewed viable arguments as a vehicle for students to justify their results, showcase their understanding, and provide a rationale for their thinking. Their vision of the role of viable arguments in school mathematics aligns to some degree with the main role of proof described by researchers as promoting understanding (Ball et al, 2002; Yackel, 2002; Hanna, 2000a). However, it is interesting that the roles of proof in promoting understanding are often presented as teacher-centered, whereas the participants' understanding of the role of viable argument was student-centered—a way for students to display their mathematical understanding rather than for the teacher to promote mathematical understanding by proving a result. Bob expressed this when he stressed the difference between the internalized knowledge a student may gain through completing an argument as opposed to the transmitted knowledge of a teacher teaching a proof.

Analysis of Construction Items

This section presents data produced as the participants solved two different construction items, by first treating each case individually and then looking at the constructions across cases. In the analysis, particular interest is given to how the participants used definitions and followed logical progressions in their construction process, how they used examples, and what features of their work allowed them to qualify their work as viable arguments. The data is again presented in sequence for David, Bob, Bridget, Sam, and Linda. Recall that the first construction item asked whether the sum of two consecutive odd numbers is divisible by four, while the second posed a question about the transitivity of divisibility.

David

There was a striking difference in how David solved the two construction items. For the first item, to show that the sum of two consecutive odd numbers is divisible by four, David first used an example to orient himself and make sure that he understood the problem correctly. He then produced a proof by using the definition of an odd number to represent two consecutive odd numbers and show that their sum is divisible by four, noting that using the definition of odd as a starting place was the obvious choice. In contrast, when facing the second construction item asking whether the transitive property holds for divisibility, David did not know how to start the problem and was ultimately unable to complete an argument to support the statement. Instead of starting with the

definition of "divides," David quickly convinced himself that the statement was true with the aid of a couple of examples but did not see a clear way to approach the problem.

David's struggle with the second item can be interpreted in two ways. First, he may not have had a clear concept image (Tall & Vinner, 1981) of divisibility and was therefore unable to produce a written or formal concept definition. When the researcher asked him directly what it meant that a divides b, he answered that it means that b is bigger than a by some factor. While he had an understanding of what it means to divide, he was unable to translate this into a usable definition. As a consequence, David did not have a well-defined starting point in the solution process and was unable to form a logical progression.

Second, David was not entirely comfortable with variables, manifested in at least three different ways in the interview. He stated on two occasions that the variables confused him. He had to convince himself that the two letters b in the problem statement "If a divides b and b divides c, then a divides c" were the same b. Finally, he expressed that he only needed to know "the other factor of b," but he was not assigning a variable to this other factor. The fact that he was uncomfortable with variables contributed to his failure to produce a viable argument for divisibility.

In the first item, David was able to recognize that it was the definition of an odd number that allowed him to construct an argument to support the statement. It would therefore be reasonable to assume that he would recognize the importance of the definition of "divides" in the second item, but this was not so. It is possible that writing an odd number as 2n+1 or 2n-1 is such a common procedure in mathematics that he did

this without reflecting on the importance of the definition as a starting point for the solution process. If this is the case, it was David's familiarity with the representation of an odd number that allowed him to solve the first problem, rather than an understanding of the argument construction process.

When attempting to answer the divisibility question, David recognized a key idea about common factors that had to make the statement true, but he did not fully grasp the implication of this key idea as he later believed he had found a counterexample. When he recognized that his counterexample was incorrect he turned back to factors, but was not able to use the key idea to produce an argument that can be classified as a general argument: "I still agree with my factoring method, but I am not seeing how that is going to help me prove it." Even though David was able to give a partial verbal explanation for why the statement must be true, he was not able to convert this into a completed general argument or a proof.

In his solution process, David employed examples as a means to ensure that he fully understood the problem statement. He stated that he liked to use an example initially because the variables "can get a little confusing," so an example made him understand what the problem was asking. He did not use examples constructively by trying to find a solution process in the examples; instead, they were solely employed as an aid to help understand the problem statement. David also used examples to search for a counterexample. The absence of a counterexample helped him gain confidence that the statement was true, which led to looking for ways to prove the statement.

Bob

For the first construction item, Bob used the definitions of odd and even to form a logical progression and prove the statement true. He used no examples in his solution process; instead he went straight for a general representation of even and odd numbers. He stated that he had essentially done the same problem last semester and knew immediately how to approach the argument. He did rate his argument as a viable argument, but claimed it was not a proof "because I did not say z is an integer. I did not go by it step by step." It is evident that he regarded proofs to adhere to requirements of both form and content, and an argument that did not adhere to the required form was not a proof, even if the content was sufficient.

To prove the transitivity of divisibility, Bob immediately recognized that "a" must be a common factor of b and c, and he used that fact to produce a proof. Before Bob completed the proof, the researcher asked if he believed the conjecture to be true or false. Bob responded: "I believe it to be true, just because it seems like if b divides a, then a is a factor of b…by definition. So then if c is divided by b, that means that c is a factor of b, and if a is a factor of b, then a is a factor of c by definition—I think," His statement was slightly inaccurate—he accidently switched the role of b and c, and he hedged his bets at the end—but it is clear that Bob grasped the key idea that solves this problem. He was furthermore able to convert this into mathematical notation to provide a proof of the statement. He did not use any examples in his argumentation process, most likely because he, unlike David, was comfortable enough with the definition of "divides" and with

variables that he did not need to resort to examples to see if the statement was true or false.

Since Bob had very little difficulty solving the two original construction tasks, he was asked to prove or disprove that the derivative of an even function is an odd function. He struggled a bit more with this task. As he explained, he initially did a couple examples of even functions and found the derivative of these functions were both even. Bob was not able to provide a correct argument because he mistook even functions and odd functions to be polynomials of even and odd degrees, respectively. Using these incorrect definitions, he was able to logically derive the result although he was not sure if he had produced a viable argument. He was clearly less comfortable with this solution and felt that he had less grasp on the material.

At the conclusion of the interview, the researcher, sensing that Bob did not know the correct definition of odd and even functions, returned to this problem, providing Bob with definitions based on symmetry. Bob then approached the problem graphically, by creating examples of even functions, but he did not attempt to write a symbolic representation of the two definitions. This may have been due to lack of time, but it may also be due to the fact that Bob was clearly not as comfortable with calculus as he was with geometry and number theory. When asked what a derivative is, Bob used his understanding of the derivative as the slope of the function to give a verbal argument explaining why the statement must be true, but he stated that he had "no idea" how to give "an argument that is not strictly words."

Bob's solution process was the same for all three construction items. He built a logical progression from the hypothesis to the conclusion of the statement using definitions, theorems, and results along the way. He recognized key ideas and was able to convert these into viable arguments and proofs. The only time Bob strayed from this process was with the calculus problem, but this may be because he had reservations about his calculus knowledge and had no experience with calculus proofs.

Bridget

Bridget's solution process was very similar to Bob's, who used definitions to build a logical progression proving each construction item. Bridget was able to prove both items in this manner, although her construction of the first solution was slightly different than the others. Instead of simply summing two consecutive numbers, Bridget stated that there needs to exist a k such that the sum of two consecutive numbers is equal to 4k. She then wrote the two consecutive odd numbers as 2n-1 and 2n+1, and proved that the k exists and that n=k.

Her solution for the second construction item employed the same solution method with the exception that Bridget checked one example prior to proving the statement. She stated that she tried the example just to make sure that the statement was true. She rated both of her solutions as viable arguments because they were general, they used definitions, they followed logically, and they started at the hypothesis and ended at the conclusion.

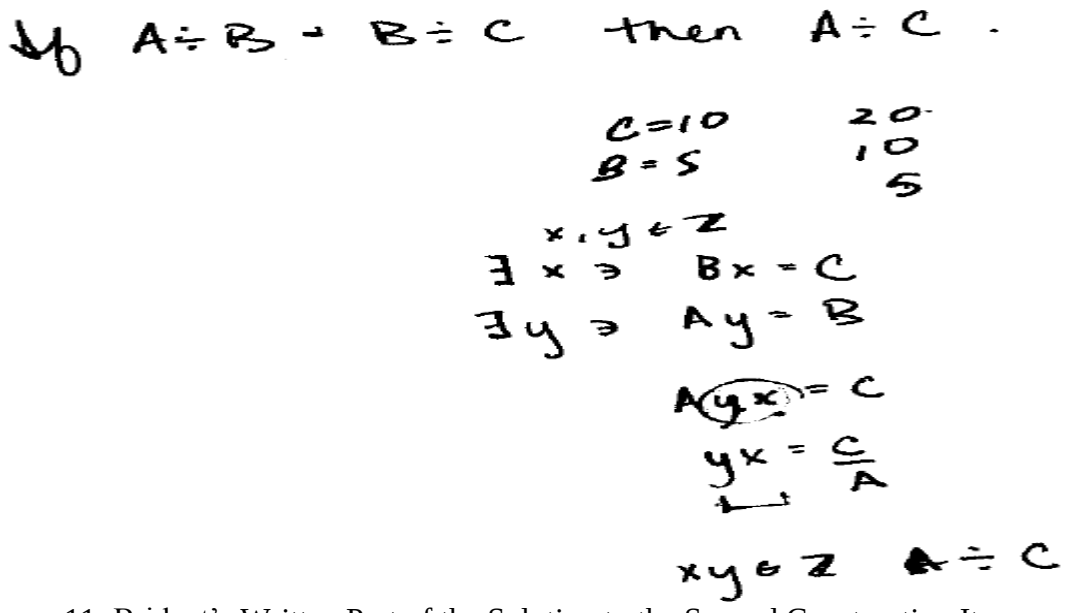


Figure 11: Bridget's Written Part of the Solution to the Second Construction Item

Bridget similarly approached the calculus item by first stating the definitions of even and odd, and she would essentially have proved the statement right away, had she taken the derivative on both sides of the equation f(x)=f(-x) instead of just on the right side of the equal sign (Figure 12). However, she was uncertain that she had used the chain rule correctly, and she did not recognize how close she was to proving the statement. She then attempted to prove the statement using the definition of the derivative as the limit of the difference quotient but was unable to complete an argument this way.

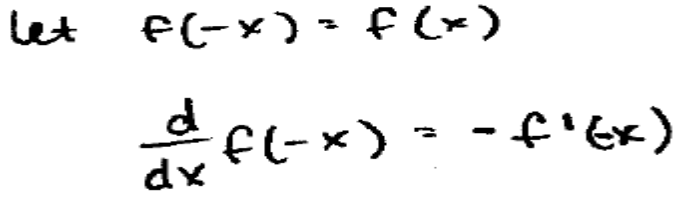


Figure 12: Bridget's Initial Solution Attempt for the Calculus Item

When asked if she thought the calculus item was true or false after her attempts to prove it, Bridget used a few examples, taking care to include the absolute value function and a trigonometric function. Failing to find a counterexample, she stated that she believed that the statement to be true but she did not know how to prove it. Bridget was clearly aware that her examples were not sufficient to verify that the statement is true. It was evident from her solutions that Bridget is a very capable student and that she knew how to construct a mathematical argument or proof.

Sam

Sam completed the first construction item without much difficulty by defining the two consecutive odd numbers as 2n+1 and 2n-1, then adding them and canceling the two ones.

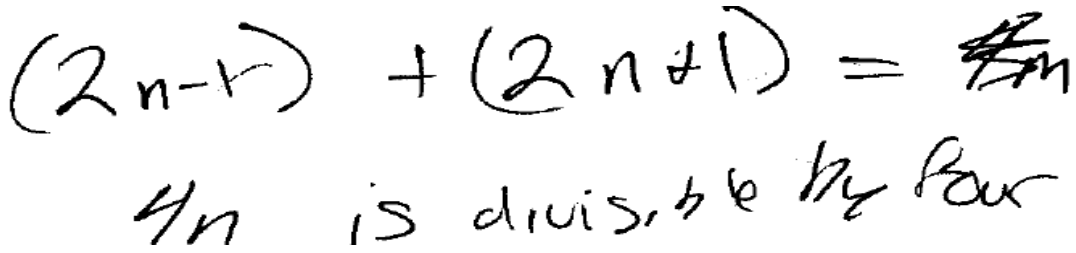


Figure 13: Sam's Written Part of the Solution to the First Construction Item.

Sam noted that he started his solution by mentally making the computation 1+3=4 since he was not sure if the statement was true or false. Then since the initial case worked, he attempted to prove the statement in general. The example served as an indicator of whether or not the statement was true and as a guide for whether he would try to prove or disprove the statement. He rated his argument as viable because of its generality, but he

also stated that it was only a rough sketch and he would have to include formal language to make it a formal proof. Specifically, the definition of an odd integer as 2n+1 would have to be formally introduced before it can be used.

For the second construction item (Figure 14), Sam completed two examples before he turned to a general approach. He used the definition of "divide" to form a logical progression and provided a general argument without much difficulty. Sam stated that the examples helped him "think about what was really happening" and helped him start to generalize his argument. He did not elaborate more on how the examples helped him with the generalization, leaving it unclear whether he was using examples to develop a foundation for his argument or simply to help him understand the meaning of the task.

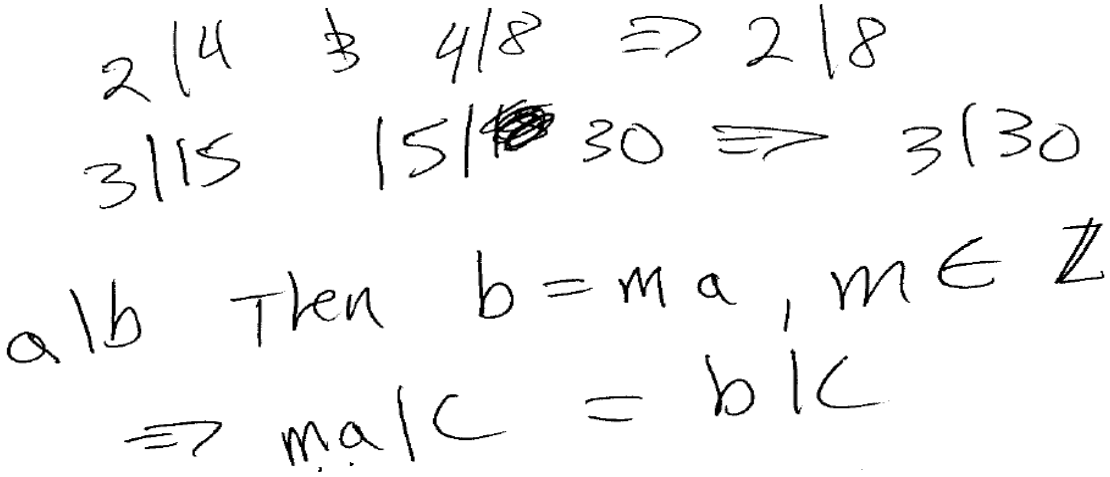


Figure 14: Sam's Written Part of the Solution to the Second Construction Item.

Sam was confident that his solution was a viable argument "as long as the definitions are right," but said he would improve it if he were to hand it in. For Sam, the argument needed to be written up formally in order to be a proof, and although he

considered his argument to be both viable and correct, the form of representation disqualified it as proof. It may seem from the written portion of the argument that Sam has not proven that a is a factor of c, but his discussion highlights that he is aware of this relationship; "b is just some placeholder for a multiple of a, so therefore c is some multiple of a." The combination of the written work with his discussion allowed the researcher to classify his argument as a proof.

It was also clear that Sam regarded whether an argument is viable or not to be dependent on audience and context. He stated that the symbolism used in his response was too advanced for a middle school classroom and that he would use the idea of common factors to explain the result in a middle school setting.

For the calculus item, Sam used the same approach as for the other two problems. He tried a couple examples of even functions to understand and visualize the problem, then attempted to prove the statement using the definitions of even and odd. The next step in his solution process demonstrated an interplay between his concept image and the formal concept definition of an even function. Sam was unable to recall that an even function is defined to be a function that satisfies $f(x)=f(-x)$, but he knew this definition exists and that it probably held the key to a proof. Consequently, Sam attempted to use a graphical representation of a function that was symmetric about the y-axis to deduce the equation for an even function but he was unsuccessful in doing so. When it became clear to the researcher that the incorrect definition hindered his solution process, the researcher provided him with the correct definitions. With this information, he was able to produce a valid proof of the statement.

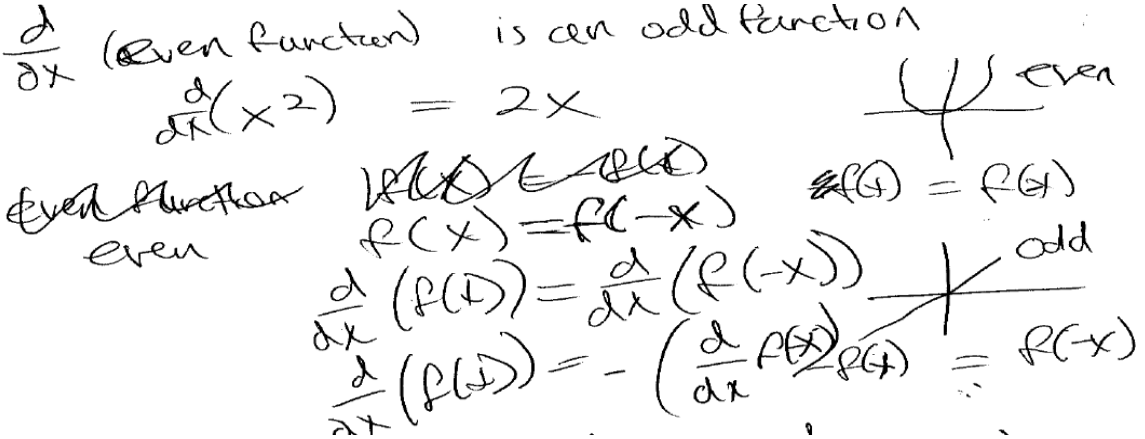


Figure 15: Sam's Written Part of the Solution to the Calculus Item.

Like Bridget and Bob, Sam had a well-developed understanding of the structure of a mathematical argument, and he was able to construct arguments starting from definitions in the hypothesis and building a logical progression to reach the conclusion. He recognized key ideas and he was able to transfer these ideas into a written argument form. He did use examples consistently to achieve an initial understanding and visualization of the problem statement, and he also used these examples to get an initial idea of how to prove the statement.

Linda

Linda expressed that she always thought of an odd number as 2n+1, so she used this fact to construct an argument showing the truth of the first construction item. She did not try any examples to convince herself that the statement was true; saying instead that she just assumed it to be true and attempting to prove it using the definition of "odd." She struggled more with the second construction item, initially claiming that it was true by the definition of the transitive property. At first she seemingly confused divisibility with

division and wanted to use fractions, but after completing a couple of examples she was able to resolve this misconception and asked if divisibility meant that one number divides the other evenly.

Linda then correctly deduced that if a divides b, that implies b=xa, and she used this as the basis for an argument. However, where Bob, Bridget, and Sam had a clear idea of how to complete the argument by recognizing that the result hinged on common factors, Linda did not have the same insight. Her solution attempt essentially consisted of symbol manipulation and she did not provide an acceptable justification of the statement. When asked if her argument was a viable argument, Linda was hesitant, but thought it was approaching a viable argument since it used generality. The majority of Linda's work on the divisibility item is shown in Figure 16.

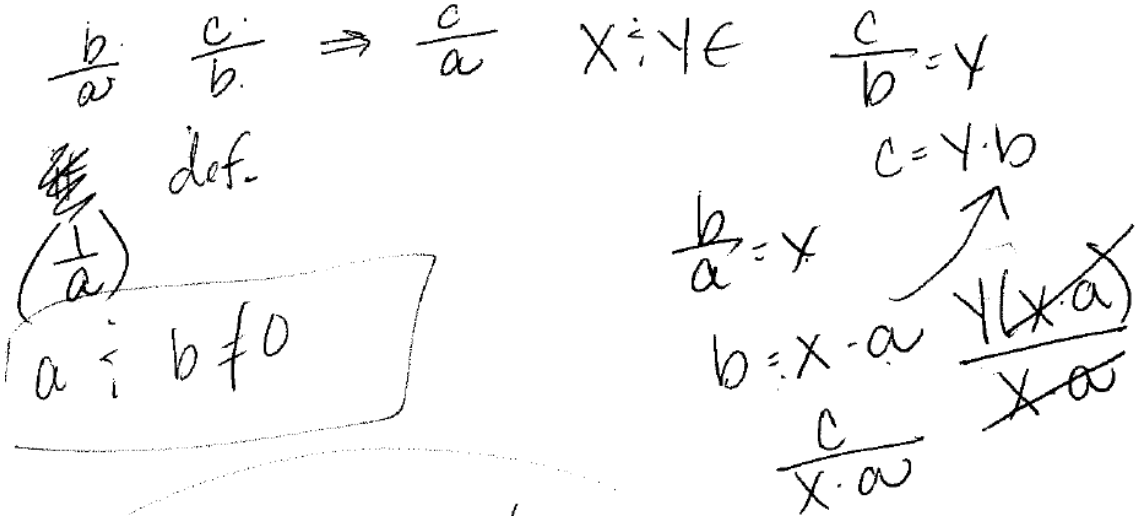


Figure 16: Linda's Written Part of the Solution to the Second Construction Item.

The cancellation in the bottom right corner of the figure suggests that Linda's argument essentially used symbol manipulation to show that $y = \frac{c}{b} = \frac{yxa}{xa} = y$. Although her exact reasoning is difficult to discern, she apparently did not recognize the fact that $\frac{c}{xa} = y$, where y is an integer, means that a divides c. Consequently, Linda did not provide a correct argument supporting the truth of the statement.

Analysis across Cases

In the ability to construct viable arguments there was a clear distinction between the three students identified as high or average-to-high achieving and the two that were identified as average or low achieving by the mathematics education faculty. Even when challenged, Bob, Bridget, and Sam relied on definitions and logical progressions to construct an argument. They displayed a clear and concise understanding of the process of argument construction, and they were able to not only see why a result must be true but also to translate that knowledge into a viable argument or a proof.

For David and Linda, the situation was somewhat different. They displayed understanding of the process behind mathematical argumentation, but when challenged they were not able to use definitions in a constructive manner. For David this manifested itself in his inability to use variables efficiently, even when he recognized the structure that explains why the statement is true. For Linda the situation was exactly the opposite: she was able to write a definition of "divides" using variables, but she did not recognize the key idea of common factors and instead resorted to symbol manipulation with no apparent plan or defined goal.

In terms of the types of arguments the participants constructed, they all valued generality as an important feature in an argument and they all attempted to prove the different statements in the construction items. Consequently, there was very little difference in how the participants approached these tasks and all the solutions to the construction items could be coded as either proofs or general arguments in relation to the 5-level rubric in Figure 8, where the difference between the two relied more on the arguments' form and representation rather than its content. It is important to note that these codes were assigned based on the structure of the arguments, rather than the correctness or the validity of the argument, although it was noted if an argument was incorrect. The specific coding of responses was based on both the written and the verbal work of the argument, so a written general argument that was supplemented by a verbal explanation of the details required for a canonical proof was coded as a proof. The codes for the responses can be found in Table 9:

Table 9: Coding for Construction Items

Participant	Item 1 – Odds	Item 2 - Divisibility	Extra Item -Calculus
David	Proof	Attempted Proof (not completed) Incomplete general argument (verbal)	N/A
Bob	Proof	Proof	Attempted Proof (incorrect) General Argument (verbal)
Bridget	Proof	Proof	Attempted Proof (incorrect)
Sam	Proof	Proof	Proof (after correct definitions provided)
Linda	Proof	Attempted Proof (incorrect)	N/A

Analysis of Evaluation Items

The following section addresses the participants' responses to the evaluation items in the task-based interview. Following the pattern of the previous sections, the five individual responses are discussed first, followed by a cross-case analysis. The evaluation items introduce a new dimension in the form of multiple samples of student work. Given the number of different student arguments to be evaluated, it is not realistic to present each participant's response to each argument, or even each assessment item. Instead, the discussion focuses on the underlying patterns found in each participant's evaluations of the nine samples of student work, with specific examples included where helpful.

Faculty Member Analysis

To ensure that the researcher-generated "student work" represented the type of arguments intended by the researcher, a senior mathematics faculty member was asked to evaluate the arguments based on the argumentation pyramid and the rubric presented in Chapter 3. The faculty member's evaluation was consistent with the researcher's analysis of the arguments also described in detail in Chapter 3. There were two separate evaluation items used in the interview, shown in Figures 17 and 18. The first were five responses to the statement: "The sum of two consecutive odd numbers is divisible by four." The second consisted of four responses to the statement: "A number is divisible by three if the sum of its digits is divisible by three."

Is the sum of two consecutive odd numbers divisible by four?

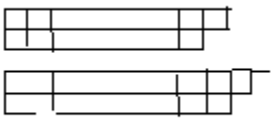
Pat's response: 1+3=4, 5+7=12
7+9=16 11+13=24
53+55=108 123+125=248
We see that all the answers are divisible by four so the sum of two consecutive odd numbers is divisible by four

Chris' response: 1+3=4
3+5=8
5+7=12
We see that the sum of two consecutive odd numbers is divisible by four.

Alex's response: Consider two consecutive odd numbers. One of these numbers must be one more than a multiple of four and the other must be one less than a multiple of four. For example, 25 is one more than a multiple of four and 27 is one less than a multiple of four. So, when you add them together, you get a multiple of four. 25+27=52, which is a multiple of four.

Taylor's response: If you add two odd numbers the answer is always even. When you add two even numbers, the answer is always a multiple of four. So, when you add two consecutive odd numbers the answer is always a multiple of four.

Ryan's response:



When these two odd numbers are added together, we see that the one box left over in the larger odd number will fill in the missing box in the smaller odd number creating a new box with four rows and no boxes left over. Since there are four rows, the number is a multiple of four. This means that the sum of two consecutive odd numbers must be a multiple of four.

Figure 17: Researcher Generated Responses for Evaluation Item 1.

A number is divisible by 3 if the sum of the digits is divisible by 3.

Reece's response:

Number	Sum of Digits	Number divisible by 3?	Sum of digits divisible by 3?
12	3	Yes	Yes
13	4	No	No
360	9	Yes	Yes
361	10	No	No
137,541	21	Yes	Yes
157,541	23	No	No

We see that the rule is true.

Gus' response:
Let a be any three digit number, with digits x,y, and z. By the place-value concept, a=100x+10y+z. This equality can be written a=(99x+x)+(9y+y)+z. By the commutative and associative properties we get a=(99x+9y)+(x+y+z). Notice that the expression 99x+9y is always divisible by 9, and therefore also by 3. Now, if the second expression, which is the sum of the number's digits, is divisible by 3, then we get that the number itself is divisible by 3.

Anna's response:
Let a be any whole number such that the sum of its digits is divisible by 3. Assuming its digits are x,y, and z, then a=xyz. Since x+y+z is divisible by three, also xyz is divisible by three. Therefore a is divisible by 3

Marina's response:
Consider 756. This number can be represented as follows: 756= 7x100+5x10+6x1. This can be rewritten 756= (7x99+7)+ (5x9+5)+6. By the commutative and associative properties, we get 756=(7x99+5x9)+(7+5+6). Notice that the expression 7x99+5x9 is always divisible by 9, and therefore also by 3. Now if the second expression which is the sum the number's digits, is divisible by 3, then we get that the number itself is divisible by 3.

Figure 18: Researcher Generated Responses for Evaluation Item 2.

It is important to note that in order to avoid bias arising from the sequencing of the evaluation items, these were handed to the participants in no particular order and the participants were asked to consider the arguments in whichever order they deemed appropriate. As a result, some of the participants considered all the evaluation items

together, whereas other participants addressed the arguments sequentially in an order they themselves decided.

David

David's main criterion for deciding whether an argument was viable was whether it was convincing or not. He stated that a viable argument needs to display the student's understanding and would be viable "if you are able to convince me that it is true." However, David's notion of what was convincing used a different standard than in his previous work. When constructing items, David initially relied on empirical evidence to convince him of the truth value of a statement before he attempted to generalize the results, but when evaluating items empirical evidence was not sufficient to convince him. To David there was a difference between an argument that is personally convincing and an argument that is generally convincing. While convincing himself might be adequate in constructing an argument, when he was evaluating he apparently adopted the criterion that the argument must be convincing to somebody else. This distinction between a private and a public domain for what constitutes conviction is reminiscent of Raman's (2003) distinction between a private and a public sphere for proof.

For David, the main difference between what is privately and publicly convincing depended on the generality of the argument and its ability to explain the result. David rated the three empirical arguments, Chris's and Pat's argument from evaluation item 1 and Reece's argument from evaluation item 2, on the basis that they had only managed to convince themselves, but were not able to convince anybody else. He regarded these arguments as starting places, as empirical evidence used initially to understand the

problem statement. For him these arguments were “conceptual” but not viable. When asked to compare Chris’s and Pat’s arguments, David saw Chris’s argument as what Balacheff (1988) terms naïve empiricism since it only contains a list of the first three cases. He saw Pat’s argument as what Balacheff (1988) terms crucial experiment since the cases are strategically picked. In David’s opinion, Pat realizes that it may not hold for large numbers, and he rated Pat’s as a better argument than Chris’s. He regarded testing other cases as Pat’s recognition of the limitations of testing only the first few and as an attempt to make the argument general. However, David clearly recognized the limitations of empirical evidence, and he rated both Pat’s and Chris’s arguments as non-viable.

For Marina’s argument using the number 471 to show the reasoning behind the result (a generic example in Balacheff’s terms), David kept going back and forth as to whether this was a viable argument or not. He initially regarded it as viable because he recognized the structure for the proof of the statement, but he also thought that the argument was incorrect since Marina does not factor a three out of the last term. However, he considered this a minor flaw that did not take away from the general structure of the argument. He also said that the CCSS do not state that an argument needs to be correct to be viable. Later David recognized that the argument is only proving the statement for the number 471, and he became uncertain as to whether or not that makes it not viable.

David’s hesitation as to whether or not Marina’s argument is viable was created by the conflict he experienced over the fact that the argument contains only one example, but it also explains why the statement is true and displays student understanding. In other

words, the argument is not general, and David recognized generality as important in a viable argument. On the other hand, he had named displaying understanding and conviction as the two cornerstones of viable argument, and Marina's argument satisfies both these. Consequently, David had a hard time deciding if the argument was viable or not. Nevertheless, he clearly recognized the merits of Marina's argument and rated the argument as better than Reece's list of six cases. In the end, David decided that the justification present in the argument trumped its lack of generality, and he regarded it as viable.

David experienced a similar conflict when he evaluated Gus's argument. Gus provides a proof of the statement, but the proof is only concerned with three-digit numbers. David recognized this as a weakness in the argument; it seemed that he initially evaluated this argument on the level of proof and rated it as non-viable. At the same time, David recognized that Gus provides a justification for his result, he has a convincing argument, and he displays a solid understanding of the result. Consequently, David wavered on his decision with this argument as well, but he ended up deciding that it is a viable argument.

The question of whether an argument needs to be correct to be viable came up on a couple of occasions in David's evaluation. For two of the arguments, Marina's response discussed above and Alex's response to the first construction item, David found what he deemed minor flaws in the arguments. For Marina's argument he stated that she should have factored out a three of the whole expression, missing the point that Marina is showing that the divisibility of the number hinges on the last term being divisible by

three. For Alex's response, David believed 1+3 is a counterexample to the statement that every odd number is one more or one less than a multiple of four, a mistake he shared with several of the other participants. However, David believed that these minor flaws did not take away from the general understanding of why the arguments are true, and both arguments were rated viable.

David was uneasy about the two arguments that have major flaws (Taylor's and Anna's respectively), and he rated them both as non-viable. For both arguments, David recognized that the logical connections between the statements were flawed and he thought that Taylor is "proving a different problem." He did not recognize that Taylor's statement "the sum of two even numbers is a multiple of four" is false. Finally, he regarded the last argument, Ryan's proof of the first evaluation item, as a viable argument. David stressed the generality of Ryan's representation and thought it was a "really good" visual argument. He further commented that Ryan's argument is not a proof since it uses a visual representation, reflecting his understanding that proofs are written arguments. A summary of David's ratings are given in Table 10.

Table 10: Summary of David's Argument Evaluations

Argument Author	Viable?	Argument Strengths	Argument Weaknesses
Chris	No		Not convincing
Pat	No	Starting place	
Ryan	Yes	General	Not a proof since visual
Taylor	No		Flawed logic
Alex	No		Flawed
Gus	Yes	Convincing, displays understanding	Only 3-digit numbers
Marina	Yes	Correct structure	One example, flawed
Anna	No		Circular logic
Reece	No		Not convincing

When asked if the arguments he rated non-viable could be made viable and if they could be improved, David focused on making them general. For the empirical arguments he stated this could be achieved through using variables. David did not recognize a pattern in Chris's argument and he did not have any suggestions for how to improve it other than representing an odd number as 2n+1.

Bob

In light of Bob's responses that a viable argument can consist of observations only, it would be logical to assume that he would rate empirical arguments as viable. However, when evaluating arguments, Bob did not accept the three arguments that represent naïve empiricism (list of examples) or crucial experiment (strategically picked examples) as viable. Instead he commented that these statements needed to be general and that the arguments, although a good start to convince the author that the statement may be true, do little to convince others. Bob shared David's sentiments that conviction has a personal and a public sphere, and that personal conviction is not enough. When questioned, Bob also stated that conviction alone is not sufficient for an argument to be viable.

The main criterion that Bob used when evaluating empirical evidence was generality. He stated that Chris's and Pat's example-based arguments for the first evaluation item needed to "recognize something and then apply it." He stated that Pat's argument went further than Chris's because it tests more examples, but he still preferred Chris's argument since it does the examples in order. Bob did not comment on the possibility of extending Chris's arguments inductively, but he mentioned three different

ways that the solution could be generalized, including his own solution to the problem. Since he did recognize several ways that the argument could be generalized, he was not concerned with how this happened as long as it happened. This was supported by Bob's comment that Reece also needed to recognize generalizing features in his argument for the second construction item, but because of the difficulty level of the argument he stated that Reece had much further to go than Chris and Pat.

Bob's responses to the two flawed arguments were very different from each other. He dismissed Anna's argument for the second conjecture as being circular and he did not find much merit in it, stating that it had a long way to go. Bob recognized that Taylor's argument had a logical flaw, but he recognized that Taylor could argue that the sum of two consecutive odd numbers will equal two times the even number between them, that is $(2n+1)+(2n-1)=2\times2n$, which is divisible by four. Bob did not recognize the false statement in Taylor's argument, deciding instead that it is true by "strong observation." It is possible that since Bob had already pointed out a logical flaw and a method to improve the argument, he was not concerned with evaluating the response in more detail. He rated both of these arguments as non-viable since they contained logical mistakes.

The remaining arguments, Alex's and Ryan's from item 1 and Gus's and Marina's from item 2, were all rated as viable arguments. Bob regarded Ryan's response favorably, labeling it a "good concrete example." He liked the visual representation, commenting that it was easy to understand and that it displayed a different type of thinking. According to Bob, "Ryan is a genius." Bob commented on the presentation of Alex's argument and wanted him to introduce variable notation to streamline the

argument. Bob stated that he initially thought the statement was false because he misread the solution, thinking that the statement was only true for every second even number. He quickly corrected this misconception and he thought Alex presented a good argument, although he was not certain that the statement that every odd number is either one more or one less than a multiple of four is common knowledge. A summary of Bob's evaluation results are shown in Table 11.

Table 11: Summary of Bob's Argument Evaluations

Argument Author	Viable?	Argument Strengths	Argument Weaknesses
Chris	No		Not general, not convincing
Pat	No		Not general, not convincing
Ryan	Yes	Concrete, visual	
Taylor	No	Contains correct idea	Flawed logic
Alex	Yes	Correct argument	No variables, needs explanation
Gus	Yes	Proof for 3-digit numbers	Only 3-digit numbers, needs additional formality
Marina	Yes	Structure of proof, 471 as variable, shows how it works	Only one example
Anna	No		Circular
Reece	No		Not general, long way to go

For Gus's and Marina's solutions to the rule for divisibility of three, Bob valued the structure of the arguments and accepted both as viable. He noticed that Gus only proves the statement for a three-digit number and to formally prove the statement he would need to fix that. He also commented that Gus should have started by saying "let x=abc be a number that is divisible by three" to make the argument clearer. Although it covers only one case, Bob recognized that Marina's argument did not depend on the number she

chose, stating that Marina could argue that "she picked an arbitrary 471" and she "is getting at variables." For Bob, Marina's argument using one example "shows how it works," which was why her argument was considered better than Reece's examples that only show the rule works for six cases.

Bridget

Bridget showed a firm understanding of the limitations of empirical evidence in that she rated both Chris's and Pat's arguments as non-viable, explaining that they were only showing examples. When asked to compare them she stated that Pat gave three more examples, "but that doesn't get him anywhere." She rated the two arguments as equally bad, and she did not offer any methods to improve them. Likewise, Reece's argument was non-viable because "he didn't consider a general case," instead of specific cases. She elaborated: "So he can list ten cases, and if there was one that he didn't get that made the conjecture false, then the conjecture would be false, even though he had ten that showed that it was right." Bridget displayed a firm understanding of both the limitations of listing examples and the role of a counterexample, explaining that the three arguments were all just lists of examples that fail to show why the result is true.

Bridget struggled a little with Ryan's argument since the visual representation of the argument seemed to confuse her. In particular, she did not quite understand what happened in the last step when the boxes are put together, and she stated that she would have to ask the student for clarification on this. She was not willing to commit to whether the argument was viable or not, stating she "thinks that it is viable" and that it was "on the right track" since it is using a general definition of an odd number and the steps

appear to be logical. She expressed, without any prompts from the researcher, that she found it difficult to offer suggestions for improving the argument as she was not quite sure what happened in the last step.

Bridget also struggled with the evaluation of Alex's argument, claiming that "it is hard to follow." She thought the argument was viable, but she was hesitant in her evaluation and seemed to have a hard time deciding whether the argument was correct or not. She expressed that she was not sure Alex is covering every case. Her hesitation stemmed from trying to decide whether Alex is saying that one number is one less than a multiple of four, and the other is one more than the same multiple of four. She mused:

> He doesn't say, but k could equal n, so then it would cover. Oh, so if they don't, okay so because he doesn't say that they are equal, we could go 4x2+1 and 4x4-1, we get 15 and 9 and they are not consecutive, so….He doesn't say that they have to be equal. One of these numbers must be one more than a multiple of four, he doesn't say that they have to be equal for them to be consecutive, so that, that's the only part that would be wrong. … but he is on the right track and I would say that it isn't completely right, but it is a good argument.

Apparently, Bridget was evaluating the converse of the claim made in the argument and as a consequence, she concluded that the argument was incorrect. However, she still thought it was a good argument, and she rated it as viable.

For Taylor's argument, Bridget immediately spotted that the statement "the sum of two even numbers is a multiple of four" is false, and she provided the counterexample that 2+4=6. This, however, did not cause her to discard the argument, and she still appeared hesitant as to whether the argument was wrong: "When you add two even numbers, the answer is always a multiple of four, is that true? No, 2+4=6, so this is wrong, I think." Bridget still found value in the argument, and in her discussion of the

response she differentiated between the form and the content of the argument. The content was incorrect, and she provided a counterexample to show that it is incorrect. Nevertheless, she found merit in the form of Taylor's argument, expressing that "he knows how to construct an argument." She commented that Taylor is starting at the hypothesis and attempts to make a logical progression to end at the conclusion. So, the argument was not in itself viable since it was incorrect, but she rated it highly in that it conforms to how a viable argument should look.

For Anna's flawed argument in support of the second construction item, Bridget recognized that the argument is circular and dismissed it as not viable and "not proving anything." She commented that Anna is starting with what she is trying to prove and ending with what she is trying to prove. Although she did not explicitly comment on the form of Anna's argument, Bridget dismissed it without much consideration and did not offer any ways to improve this argument.

Bridget rated Gus's argument as viable because it is starting with the hypothesis and following a logical sequence to show that the statement is true. She also commented on the generality of using variables, and briefly mentioned that Gus is not saying anything about numbers that are not three digits. Bridget further stated that Marina's argument would be viable if it was more general. Like Bob, Bridget valued Marina's single example more than Reece's six examples, because Marina is able to show why the result works.

In general, Bridget's evaluation was hesitant and she seemed to be struggling with making a distinction between evaluating the responses as viable arguments versus

evaluating them as proof. She recognized that Marina's response explains why the divisibility rule for three holds, but she hesitated to call it viable because it is only one example. Similarly, she initially expressed hesitation about Gus's response because it does not take anything other than three-digit numbers into account. Bridget also made several mistakes in her evaluation of the argument, which is surprising in light of how easily she was able to construct arguments herself. Table 12 provides a summary of Bridget's evaluations.

Table 12: Summary of Bridget's Argument Evaluations

Argument Author	Viable?	Argument Strengths	Argument Weaknesses
Chris	No	Only showing examples	Only showing examples
Pat	No		Only showing examples
Ryan	Yes (hesitantly)	On the right track, generality, uses definition	Confusing
Taylor	No	Good form and structure	False statement
Alex	Yes		Hard to follow
Gus	Yes	Logical, general	Only 3-digit numbers
Marina	No	Logical	Not general
Anna	No		Circular
Reece	No		Not general, list of examples

Sam

To a large extent, Sam's evaluation was similar to that of the other participants. He rated Gus's argument as viable, even though it is only concerned with three-digit numbers. He recognized and pointed out the flaws in both Taylor's and Anna's arguments, noting that the second line in Taylor's argument is false and that the steps are not logically connected. Likewise, Anna's response was identified as circular since Anna "used the conjecture to prove that the conjecture is true." Ryan's argument was described

as a visual and creative argument that has a nice logical outline. Like Bridget, Sam stated that he would like to see a bit more explanation in Ryan's argument about how the last box shows the divisibility by four, but overall the argument was deemed viable. A summary of Sam's responses appear in Table 13.

Table 13: Summary of Sam's Argument Evaluations

Argument Author	Viable?	Argument Strengths	Argument Weaknesses
Chris	No	Good start	Only examples, no understanding
Pat	No	Good start	Only examples, no understanding
Ryan	Yes	General, structured	Confusing, need clarification
Taylor	No		Logical flaw, false statement
Alex	Yes	Logical, concise, visual, nice outline	
Gus	Yes		Only 3-digit numbers
Marina	No	Starts to show generalization, good start	Only one example
Anna	No		Circular
Reece	No		Only examples, no understanding

Sam rated Alex's argument as viable, as it was logical and concise. He was hesitant at first about Alex's assertion that if you have two consecutive odd numbers, then one is one more than a multiple of four and the other is one less than a multiple of four. Sam spent some time considering this statement, then claimed that it is true because he could not find a counterexample. At a personal level, Sam convinced himself that the conjecture was true on the absence of counterexamples.

In stark contrast to this was Sam's evaluation of the empirical arguments among the two items. Although on a personal level, the absence of a counterexample seemed to

convince Sam, he displayed a clear understanding of the limitations of empirical evidence, stating that "We could have a 100,000 particular cases and there is one case where it doesn't hold true and the conjecture is false." Consequently, Sam rated both Chris's and Pat's arguments as not viable since they do not demonstrate any understanding of why the statement is true. The following dialogue highlights Sam's understanding of the use of empirical evidence in proof:

> Sam: Well, honestly I would say that [Chris and Pat] haven't demonstrated any different level of understanding from each other. I mean, this one [Pat] it is more cases shown, but this person [Chris] could have worked out the same if he just picked some cases where it was true. This person just picked out more, he has still not done anything to prove it for all cases. It is just looking at a very select few. So I would say that overall none of these arguments, none of them are any better than the other.
>
> Interviewer: Okay, can any of them be fixed….Do any of them have any merit whatsoever?
>
> Sam: They do for their particular cases, and this is definitely, both of them are a good start. From looking at it, you definitely want to run through a few in your head trying to see if you can find a counterexample or if they had found a counterexample then they would have just been able to state that and then done. Uhm, the next step for both of these would be to take it to a more general case like over here, and then it would be perfectly fine. You would have some extra evidence showing that they had some cases to show that it was true to go with the general argument.
>
> Interviewer: Could you make any of them more general, if that is what would make it viable?
>
> Sam: So they would have to bring in and introduce some sort of variable or bring in some sort of visual like this argument here [Ryan] to where they can talk of any two consecutive integers. It's tough to say how they could go about it, because there are a lot of ways that they could and still have it be a viable argument. Uhm, …

It is clear from the this exchange that Sam fully recognized the limitations of empirical evidence, and that he did not rate crucial experiments as better than naïve empiricism. Both arguments were a good start for the initial conviction that the statement is true, but both arguments failed to generalize and show that it is true for all cases regardless of whether they tested three or six cases. It is also evident that Sam believed generality is achieved through use of variables. "[If] I was to give him [Chris] a comment it would be what about for integers, or for odd consecutive integers 2n-1 and 2n+1, I guess. Just to kind of get him thinking of the form that might help him prove it for all cases." Sam did not offer any other way to extend the argument.

To see if he recognized that Chris's argument has the start of a pattern that can be explained, the researcher asked several follow-up questions about possible extensions.

> Interviewer: What if Chris here had said, 'we see from the pattern that the sum of'….If he includes those three words, would you have treated it differently?
>
> Sam: I still feel that yes he may have developed a pattern here, but there is nothing that is saying that once I get higher and higher into it that maybe that pattern doesn't hold. There is nothing governing the nth case if you will…So maybe, I definitely acknowledge that maybe he was giving some thought to a pattern but still there is nothing to contribute to what that pattern was.

Sam did not explicitly recognize a pattern in Chris's work, and he dismissed the argument as naïve empiricism. His dismissal of empirical evidence also extended to generic examples. When evaluating Marina's argument, Sam focused on the fact that she has only proven the statement for one particular number and that the argument needs to be generalized. In doing so, Sam was not recognizing the general form of the specific

argument, even when directly asked if there was a difference between Marina's generic example and Reece's list of six examples.

> Interviewer: So if you look at the six examples, and the one example….Any one better than the other?
>
> Sam: Well, they both have their strengths in that this person definitely sees that okay, if I take the 100s place I can multiply that by a 100, add that to the tens place, multiply that by ten, add the placeholder for the ones place and add that up. So they got a good start here, so that's good—and then this one it goes into more cases where he shows where they are divisible by three, then the sum, the sum of the digits is also divisible by three and then likewise he showed that if the digit sum is not divisible by three, then the number is not divisible by three, so I would say that they are probably on equal setting. That they are just, they are different if that makes sense. They both were considering different ways to go about it and they both have their strengths and they both have their shortcomings I suppose.

At this stage it seemed to the researcher that Sam was evaluating proof and not viable arguments. To investigate this further and to see if Sam would change his answer when reminded that there may be a difference between proof and viable argument, the researcher asked about the role of viable arguments:

> Interviewer: Well, I just have one question that I would like to repeat from last time…what is the role and the purpose of viable arguments in school mathematics? Why do we care?
>
> Sam: I think it is probably the most important part of math, is to be able to justify something that you do or everything that you do, at least to some extent—to say here is why I know what I know, and here is why I do what I am doing. It's kind of where it is coming from. It doesn't necessarily have to be a formal proof, but at least the thought behind it, saying 'I know this to be true so I can use this result or this process.' If that makes any sense.
>
> Interviewer: So, in light of what you just said, justifying….Any difference between these two?
>
> Sam: As far as like justifying?

> Interviewer: No, as far as providing….The role of a viable argument is to provide a justification for the result that you have, right?
>
> Sam: Right.
>
> Interviewer: Is there any difference between these two?
>
> Sam: Well I suppose the only difference I can think of is that…this one just provided one case justification for one case, whereas this one provided justification for six cases where it held. In that regard, this one [Reece's] would be better, though I feel that this one [Marina's] start to show more of the generalization, kinda getting to that point.
>
> Interviewer: Would that make it viable, that it is getting to that point?
>
> Sam: Not as it is, but if you look at the good one here [Gus], the form is fairly similar, even though it is just a particular case, so (inaudible) take it to the next level. I know this person [Marina] is much closer than this person [Reece]. All this person [Reece] has done is showed me six cases where it is true. So, in this regard this [Reece's] is better, but as far as taking it to the next level, proving it for all cases, this person [Marina] has a much better start I would think.

Sam's response here differed considerably from the responses provided by Bob, Bridget, and David, who recognized the structure of Marina's argument as sufficient to be labeled a viable argument. Sam also recognized a difference in format in Marina's argument, but he discarded the argument as only proving one case although it was a better start towards a viable argument than Reece's list. When defining viable arguments, Sam stated that there is a difference between proofs and viable arguments, but when evaluating arguments, he was essentially evaluating the arguments on the level of proof. Consequently, Marina's argument is not viable since it is not general.

Linda

Linda was very hesitant in her statements about the evaluation tasks, and she commented that she was on the fence on several of the student responses. Her evaluations were delivered without conviction and there were several long pauses within her discussion. She was consistently hedging her opinions on whether the arguments were viable or not, as witnessed in the following quote from her evaluation of Alex's argument from the first evaluation task.

> Oh yes, I guess that is true, yeah. …(mumbles) but does that…yeah I guess it does. Uhm, you can see that they have an understanding of it and that it is true. Uhm, I don't know if it is a viable argument, I mean I know that if I was actually grading this, I would have liked to have seen maybe like a…4n-1 and 4n+1, I mean written out a little bit more, just to translate their understanding and their thought process on this a little bit better. Uhm…yeah.

When asked directly whether the response was viable, Linda answered:

> I am really on the fence about this one, I don't know, uhm….Well, with a little bit more explanation, yeah. Maybe if they were like talking through it with you. …I just like to know…uhm, with the 4n-1 and 4n+1 how and why did they choose those and were they (sic) prior knowledge.

In addition to her inability to decide if the argument is viable or not, Linda, in using the same variable for both 4n+1 and 4n-1, did not recognize that the same multiple of four is not necessarily used to create the two odd numbers, although the example given in the argument is 27 (one less than 28) and 25 (one more than 24), which uses different multiples of four. She made a similar mistake in her evaluation of Taylor's flawed argument when she did not recognize that it contained a false statement: "The sum of two

even numbers is a multiple of four." However, she recognized the second mistake in Taylor's argument, the break in logic, and rated it non-viable.

Linda recognized the limitations of empirical evidence by noting that all three lists of examples failed to account for a general case. She still was indecisive when asked if this made the arguments non-viable, indicating that these arguments showed an understanding that the statement was true and that they were "good arguments," but they failed to provide a general argument. She interpreted Pat's argument and examples as "showing some understanding that it continues," since the argument addresses cases beyond the first few examples. She did point out that the arguments were not proofs, but she was not sure if they were viable arguments although she decided in the end that they were not.

Linda consistently addressed whether the arguments displayed understanding in her evaluations of whether they were viable or not. She rated Marina's example as better than the other arguments that used examples, based on the fact that it outlined why the statement is true. When asked if this affected whether she considered the argument viable or not, she answered the following:

> Yeah. Because it shows, I mean, …it shows an understanding, and if you are able to explain the why, then it is an argument that can be conveyed to somebody else, rather than just having this knowledge and putting it on paper with certain numbers. You can't necessarily convey your understanding, or why it is true or false.

She did, however, note that Marina's argument addressed only one number, and added that Marina should have included a statement expressing why this would work for all cases. She did not mention that this would require variables. Finally, Linda was not

entirely happy with Gus's argument, expressing that she felt he leaves something out in the end, but it showed his understanding and she rated it viable. A summary of Linda's argument evaluation is provided in Table 14.

Table 14: Summary of Linda's Argument Evaluations

Argument Author	Viable?	Argument Strengths	Argument Weaknesses
Chris	No	Shows understanding	Not general
Pat	No	Shows understanding	Not general
Ryan	Yes	General, structured, explains	Not a formal proof
Taylor	No		Logical flaw
Alex	Yes (hesitantly)		Needs variables
Gus	Yes	Shows understanding	Only 3-digit numbers
Marina	Yes	Shows understanding	
Anna	No		Logical flaw
Reece	No	Shows understanding	Not general

Analysis across Cases

When evaluating arguments in terms of viability, the participants valued explanatory power. The participants all commented on the insufficiency of using empirical lists in showing why a statement over an infinite set is true for all cases, and they all rated the three empirical lists as non-viable. This means that issues of generality were important in the participants' decision to rate arguments as viable or non-viable, and only general arguments or arguments that contained generalizing features were recognized as viable. That is, only arguments from the top three tiers of the argumentation framework in Figure 20 were considered viable.

There was some disagreement among the participants of whether or not Pat's response that included some larger numbers was a better argument that Chris's response,

which only included the three first cases. They all interpreted Pat's response as a recognition that the result may not hold for all cases, but they also recognized that it is not convincing. All the participants except Sam valued the structure in Marina's generic example argument and deemed this a better argument than the empirical lists. None of the participants commented on a possibility to extend a pattern in Chris's argument containing three initial and sequential cases.

Similarly, all the participants recognized the flaws in the two flawed arguments, but only one of the participants provided a counterexample to the false statement that the sum of two even numbers is a multiple of four. Suggestions for how the arguments can be improved were only given sparingly, and all the participants focused on making responses general in their suggested improvements.

Conclusion

The results of the research reveal that there are obviously many differences among the five subjects and their understanding of viable argument, but there are also striking similarities in some areas. On the whole, the participants commented that viable arguments allow students to communicate mathematically and justify their results without having to adhere to the strict representation requirements of a proof. However, the participants also revealed individual differences in what they regarded as a viable argument.

Taken as a whole, the data gathered from the five task-based interviews offer insight into how pre-service teachers, and learners in general, perceive and judge viable

arguments. Chapter 5 makes explicit these insights and offers ideas for using this information to prepare teachers for engaging students in constructing viable arguments and critiquing the reasoning of others.

CHAPTER 5

CONCLUSIONS

Introduction

The purpose of this chapter is to present conclusions based on the research findings in Chapter 4. The conclusions are organized in response to the three different research questions, but other results and conclusions that emerged from the data will also be outlined. The final sections of the chapter will address implications of the findings, limitations of the study, and directions for future research.

Overall Research Conclusions

The results in Chapter 4 suggest that pre-service teachers' understanding of the term "viable argument" is complex. The pre-service teachers in this study often used different criteria for what they rated as a viable argument across the three different domains of defining a viable argument, constructing a viable argument, and evaluating viable arguments. It was necessary to explore the participants' thinking in all three contexts to build a complete picture of how they interpret, use, and judge viable arguments. Useful conclusions can be drawn regarding each of these three domains and are reported below.

Defining and Constructing

When *defining* a viable argument, the participants used criteria that closely aligned with the description of viable arguments in the Common Core State Standards. They emphasized use of definitions and logical progression of statements. When asked to *construct* viable arguments, they approached this question as if they were asked to prove statements, but without the formal representation that a mathematical proof requires. They did not regard the arguments they gave as proof since they lacked formalistic features that the pre-service teachers deemed important in proof, such as initial statements like "Let $n \epsilon Z$." The term "viable argument" escapes these requirements, and the pre-service teachers in this study typically did not include them in their solutions. The participants' constructions of arguments were influenced by several factors—mathematical understanding and proficiency, familiarity with the solution pattern, and confidence in their own mathematical understanding—but overall they all constructed arguments by starting with a definition and building a logical progression to reach the desired conclusion. They prefered to conduct this process on an abstract level, without the use of illustrative features such as examples or drawings.

Viable or Not?

Several of the pre-service teachers in this study struggled to *evaluate* certain arguments as viable or not. This struggle was not due to a lack of understanding of the argument or a deeper conceptual misunderstanding of argumentation. Rather, they often correctly identified features of arguments that were correct or incorrect, and they often suggested improvements based on this understanding. It is therefore likely that their

hesitation about viability stems from confusion about the term "viable argument" itself. The confusion is understandable in that viable argument is by design a fluid term that defies definition. To ask for a definition of viable argument or ask about the viability of certain arguments is to impose a category and structure on a term that defies such categorization. In a sense, asking to categorize arguments as viable or non-viable is making the same mistake as asking if an argument is a proof or not based on some external classification of what constitutes a proof.

If the goal is to uncover mathematical thinking, an argument must be evaluated on its own merits and shortcomings. Whether or not it adheres to a structure that would allow it to be categorized as viable argument is context-dependent and of secondary importance. "Viable argument" is designed to be an umbrella term that covers mathematical arguments of a wide range of different sophistication levels, determined by grade appropriateness and individual mathematical background. Without a clear understanding of these contexts, it is difficult to classify whether arguments are viable. Regardless of this seeming contradiction it was important to ask the question in order to understand what pre-service teachers value in mathematical argumentation as they are called by the Common Core State Standards to make argumentation a core practice in their classrooms. This will not be an easy task, especially seen in the light of the literature that uniformly describes the difficulties associated with proof and argumentation. The next section presents a finding that illuminates the complex relationship between proofs and viable arguments.

Constructing vs Evaluating Arguments

Overall, the participants appeared better prepared to effectively construct arguments than to evaluate arguments constructed by others. This manifested itself in both direct and indirect ways. More than one participant was unable to provide a counterexample when evaluating a student's statement "the sum of two even integers is a multiple of four" even though he or she sensed it was incorrect. Several other evaluation errors and oversights were made in the course of the interviews. Whereas there was a clear distinction in proficiency of constructing arguments between the students initially rated as high achieving and the participants initially rated as average or low achieving, a similar distinction did not emerge for the evaluation items. With the exception of Bob, all the participants either made errors in evaluation of an argument, or evaluated arguments in a manner that contradicted their definitions of viable argument.

Perhaps most noteworthy is Sam's inability to recognize the explanatory power of Marina's generic example for the divisibility by three task when he struggled to accept her argument as a widely applicable example. Bridget displayed a different set of misperceptions when evaluating the work of Alex, Ryan, and Taylor (described in detail in Chapter 4). This is especially intriguing when considered in conjunction with the ease with which both Sam and Bridget constructed their own arguments. This seems to suggest that proficiency in constructing arguments does not necessarily correlate to proficiency in evaluating arguments.

The complex relationship between constructing and evaluating arguments is an interesting one that warrants a more detailed study. It is possible that the ability to

construct an argument requires a different mathematical skill-set than the ability to evaluate a mathematical argument. For example, the constructor of an argument is actively trying to convey mathematical understanding whereas an argument evaluator must be able to discern mathematical understanding in the arguments of others. Future studies could investigate this relationship and its contributing factors. It is also worth exploring whether students' success in constructing and evaluating arguments reflects a tendency to emphasize construction over evaluation in proof-based courses. Pre-service teachers in particular need both skill sets as they not only construct arguments for students but also evaluate students' work and, further, engage students in critiquing each other's arguments.

Viable Argument vs Proof

The participants regarded a viable argument as similar to proof, yet the rigor and representation requirements are less stringent in a viable argument than in a proof. According to the participants, proof adheres to strict rules of representation where qualifiers are included, reasoning is logical, and theorems are cited. For a viable argument these requirements are not as formal. In this respect, the participants seem to regard viable arguments as a pre-formal proof (Blum & Kirsch, 1991), and they see the main role of a viable argument to explain and justify. This interpretation of viable arguments and proofs aligns with the view that proof has a personal sphere and a public sphere (Raman, 2003), where the personal sphere embraces personal conviction and understanding while the public sphere embraces the communal aspects of proof including the formal write-up into a well-defined format. The terminology of viable arguments

allowed the pre-service teachers to value justifications and the demonstration of personal understanding without having to conform to the requirements of formality inherent in proof.

Empirical vs Deductive Justifications

Another finding from this study provides a new perspective on how novices use and understand empirical evidence. All the participants displayed a firm understanding of the limitations of empirical evidence in proving a mathematical statement over an infinite set. Without exception, all the participants rated empirical lists as non-viable. Whether these empirical lists consisted of strategically picked examples or not did not affect this rating. Naïve empiricism, a list of examples, or crucial experiments, strategically chosen examples, were both regarded as non-viable. Participants considered these arguments to be good starting points to convince oneself that a statement is true, but insufficient to convince somebody else. This failure of examples to convince others, coupled with the fact that lists of specific examples do not generalize the result to cover all possible examples, were the two reasons that empirical evidence was disregarded as a viable argument. The finding that the participants in this study are well aware of the limitations of empirical evidence in mathematics differs from findings in previous studies, and it indicates that the pre-service teachers enrolled in this teaching program understand this distinction.

The conclusion that the participants understand the limitations of empirical evidence is based on an analysis across all three domains of this research. For example, Bob expressed in his definition that viable arguments are based on observations with an

explanation. Taken alone, this response can be interpreted that Bob does not fully understand the limitations of empirical evidence. But when seen in conjunction with his construction and evaluation of arguments, it is evident that Bob emphasizes explanation over observation and that he is fully aware of the limitations of empirical evidence. In a similar manner, Sam convinces himself that the statement "For two consecutive odd numbers one has to be one less than a multiple of four, and the other one more than a multiple of four" is true since he cannot construct a counterexample. This may be interpreted as a naïve approach, but seen in context it is evident that although Sam may well convince himself through empirical reasoning, he would not use this approach to convince somebody else. Other participants expressed similar sentiments that examples can be used for personal conviction, but have no place in argumentation designed to convince others.

Conclusions Based on Research Questions

The following are conclusions directly pertaining to the three research questions. These conclusions emerged from the results described in Chapter 4.

Conclusions for Research Question 1

Research Question 1. How do pre-service teachers define the notion of "viable argument"? Is their notion consistent with viable argument as defined by the Common Core State Standards?

In this study, pre-service teachers' definitions of viable argument were concerned with both the characteristics of the mathematical argument and with the effect and outcomes of the arguments. A framework depicting how these different aspects interact is shown in Figure 19. The characteristic features of viable arguments identified by the participants are many of the same characteristics often associated with proof. A viable argument needs to employ a mathematical approach that uses definitions and a logical progression starting at the hypothesis to reach the conclusion. Additionally, to be viable an argument needs to address features of generality that explain the argument for all possible cases.

The participants also indicated that viable arguments are context dependent. What is a viable argument may depend on the form and the representation of the argument. Additionally, and in contrast to proof, the content of the argument must be age-appropriate for the argument to be considered viable. Consequently, a viable argument in a 5^{th}-grade classroom is different from a viable argument in high school, and what is considered viable for one audience is not necessarily viable for another.

Finally, the data suggest that viable arguments are also defined by their outcomes. In order to be a viable argument, it is vital that the argument displays mathematical reasoning and thinking and that it explains why the result is true as opposed to just stating that it is true. A viable argument should be understandable and be able to communicate the mathematical result to others. The viable argument serves as a precursor to formal proof, and in many settings a viable argument is very similar to a proof but with less formal requirements.

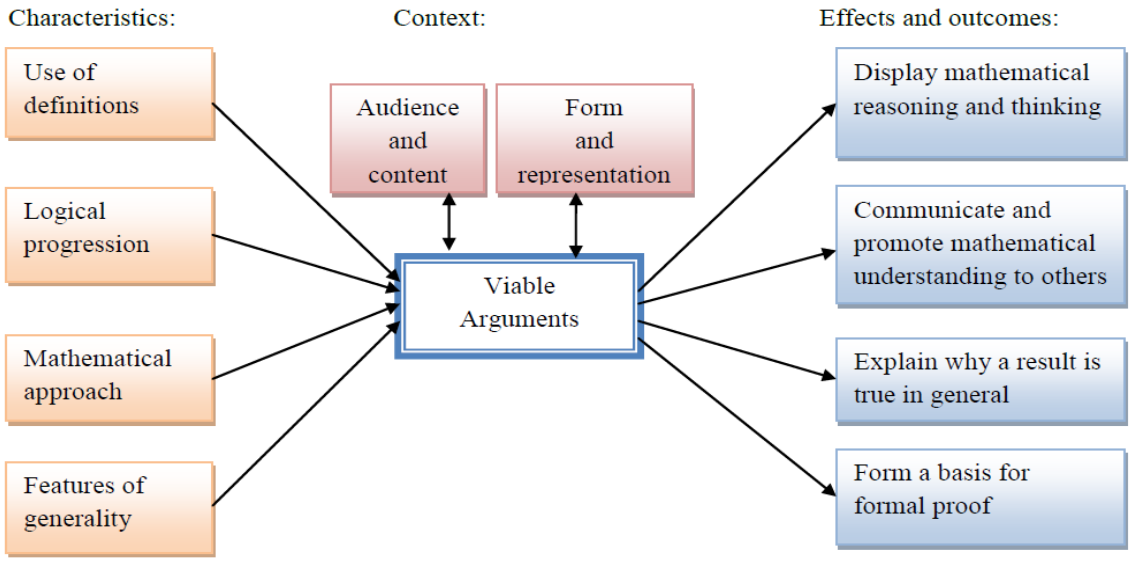


Figure 19: Diagram of Participants' Definition of Viable Arguments.

This definition of viable argument aligns closely with the description of viable arguments provided by the Common Core State Standards:

> Mathematically proficient students understand and use stated assumptions, definitions, and previously established results in constructing arguments. They make conjectures and build a logical progression of statements to explore the truth of their conjectures. They are able to analyze situations by breaking them into cases, and can recognize and use counterexamples. They justify their conclusions, communicate them to others, and respond to the arguments of others. They reason inductively about data, making plausible arguments that take into account the context from which the data arose. Mathematically proficient students are also able to compare the effectiveness of two plausible arguments, distinguish correct logic or reasoning from that which is flawed, and—if there is a flaw in an argument—explain what it is. Elementary students can construct arguments using concrete referents such as objects, drawings, diagrams, and actions. Such arguments can make sense and be correct, even though they are not generalized or made formal until later grades. Later, students learn to determine domains to which an argument applies. Students at all grades can listen or read the arguments of others, decide whether they make sense, and ask useful questions to clarify or improve the arguments. (CCSS 2011, pp.6-7)

It is evident that the pre-service teachers in this study emphasized the same elements that are described in the Common Core State Standards, and it is worth noting that they identified these characteristics prior to seeing the above description. The pre-service teachers seemed to recognize the construction and evaluation of viable arguments as developing a practice where students use a mathematical approach to effectively communicate and justify their results.

Conclusions for Research Question 2

Research Question 2. What strategies do pre-service teachers employ when asked to construct viable arguments?

2a: Do they build a logical progression of statements to explore the truth of conjectures?

2b: Do they use examples constructively in their arguments?

Logical Progressions. In general the pre-service teachers in this study attempted to create proofs when asked to construct a viable argument. They started with definitions and built a logical sequence of statements to reach the conclusion. It is evident that the pre-service teachers had a strong understanding of the process of constructing a mathematical argument, whether a proof or less formal than a proof. The participants rated their own arguments as viable arguments since they followed logical progressions and did justify the result, but they generally did not regard their arguments as proof because the arguments were not presented in the format required for a proof.

Examples. The participants used examples sparingly in their arguments, and when they did it was to ensure that they understood the problem statement correctly or to look for a counterexample. None of the pre-service teachers in the study said they used examples to generate key ideas or to see an underlying structure that could lead to a proof. Furthermore, the pre-service teachers relied on deductive arguments and none of them presented empirical arguments as a solution to the construction tasks. On the contrary, all of them commented on the limitations of examples in explaining mathematical results that involve an infinite amount of cases.

Several of the participants seemed to differentiate between a personal and a public sphere of conviction. In the personal sphere of conviction they used examples for three particular purposes. First, examples allowed them to understand the problem and visualize the problem statement with numerical values rather than variables. Second, examples helped them look for ways to prove a statement false. By noting that only one counterexample is needed to prove a statement false, the participants expressed an understanding of the implications of a counterexample in mathematics. Third, the absence of an immediate counterexample allowed the participants to hypothesize that the statement was true and pursue a way to show that it is true. It should be stressed that the participants were aware that the absence of a counterexample does not prove a mathematical statement and is not sufficient to generate a conviction in the public sphere. They applied a personal strategy to decide if a proof rather than a counterexample should be pursued, with the conclusion firmly anchored in the personal sphere.

There were three cases where the pre-service teacher failed to create a written or a correct argument for one of the construction items. David was unable to provide a definition of "divides" to construct an argument for the first construction item and, on the same problem, Linda was unable to come up with an idea to solve the problem. Although she was able to define "divides," she did not see how this definition helped the argument, and she resorted to symbol manipulation that did not create a valid solution. For the calculus item, Bob used an incorrect definition of an even function and consequently failed to construct a correct argument for the statement. When given the actual definition of an even function, Bob was able to explain why the statement had to be true, but stated that he was not sure how to prove that formally. The failure to produce arguments in support of statements was thus caused by two different factors: the inability to provide a correct and usable definition, and the inability to produce a logical progression.

Conclusions for Research Question 3

Research Question 3: What characteristics do pre-service teachers assess to critique the reasoning of others?

3a: Do they distinguish between correct logic and logic that is flawed?

3b: Do they look for ways to clarify or improve the arguments of others?

3c: Do they recognize and value constructive uses of examples?

The characteristic features of the participants' approach to argument evaluation was perhaps most clearly on display in their evaluation of Gus's response on the second construction item, where he proved a statement true for three-digit numbers. Gus's proof

gives a clear structure for how a proof could be constructed for this statement, and it also explains why the statement is true. However, all of the participants mentioned that Gus only proves his result for a three-digit number; hence the argument is not general and it is not a proof. They all recognized that Gus does provide the general structure and they all rated it viable as a result.

This example illustrates that the pre-service teachers tended to evaluate the responses on the level of proof first, where features of generality are of great importance. They then applied measures to determine if the argument contained the features that would make it viable, the most important being explanatory power and ability to convince. There were some discrepancies as what was regarded as convincing. Sam accepted an argument constructed with the aid of dynamic software as viable on the merits of showing the result for an infinite amount of cases. Bridget did not accept the same argument because it has great visual power but no explanatory power. In other words, it shows that the statement is true, but it does not explain why the statement is true.

Flawed Logic. The participants recognized circular logic and otherwise flawed logic in the arguments; however, some of the participants did not notice that the statement "The sum of two even integers is a multiple of four" is false. Only one of the participants gave a counterexample to this statement; one said that it was false without providing a counterexample; one questioned its validity but had spotted a logical flaw elsewhere and did not investigate further; and two did not catch that the statement is false. It is possible that the failure to seek a counterexample to this statement stems from

lack of experience with incorrect statements, but it is also possible that the obvious logical break in the next line of the argument took focus away from this statement. The latter interpretation is supported by the fact several of the participants incorrectly gave 1 and 3 as a counterexample to the statement that "For two consecutive odd integers, one is one more than a multiple of four, and the other is one less than a multiple of four," indicating that they recognized counterexamples as a means to prove statements incorrect.

Improvements. On several occasions the participants offered suggestions for how to improve or clarify the arguments provided in the evaluation items. The suggestions for improvement were sometimes for content in that the argument needed to incorporate generalizing features, usually in terms of using variables. In other cases the pre-service teachers offered suggestions to improve the form or representation of an argument. They also expressed concerns about what could be considered previous results and on a couple of occasions they wanted additional backing for statements they did not consider common knowledge.

Examples. The participants recognized constructive use of examples in that they recognized that Marina's solution of the second evaluation item gave the blueprint for how to prove the conjecture. All but one of the pre-service teachers rated this as a viable argument, and the participant that did not rate it viable recognized that the example was better than an empirical list. However, the participants all saw Chris's use of the three initial cases for the first construction item as what Balacheff (1988) labeled naïve

empiricism. None of them commented on a possible inductive explanation of why the pattern continues; instead all suggested improvements to this argument that involved writing an odd number as an algebraic expression, 2n+1. This result is somewhat surprising given the focus on induction proof in Higher Mathematics for Secondary Teachers, and the focus on pattern recognition in teacher education.

Implications for Teacher Education

A case study is not sufficient to make generalizations, but the findings in this study indicate that a teacher preparation program that includes a proof foundation course and at least two other subject matter courses that highlight proof positively affects pre-service teachers' construction and evaluation of arguments. The participants in this study did not display any of the naïve behavior prevalent in the research literature on proof and reasoning; on the contrary they all displayed a sound understanding of mathematical proof, reasoning, and argumentation.

Although the participants demonstrated knowledge of mathematical argumentation that is sufficient to implement the CCSS recommendations, the findings indicate that some aspects of mathematical argumentation could receive more attention. The first of these is more focus on evaluation of mathematical argumentation, particularly incorrect mathematical argumentation. Some of the participants had difficulty pinpointing what was wrong in the flawed arguments, and overall they were less confident in their answers to the two flawed responses than in the other responses from the evaluation items. Furthermore, none of the participants commented on the possibility of extending a

pattern in Chris' argument listing the three first cases to show that the sum of two consecutive odd numbers is even, instead suggesting improvements based on introducing variables. These findings are perhaps a little disconcerting as the ability to evaluate arguments is crucial in fulfilling the requirement that students critique each others' reasoning as well as an important part of assessing student work. Additionally, it is important that the evaluation is done on the students' terms and not on the teachers.' Improvements to arguments need to be suggested based on the merits of the students' work and not on the teachers' preferred solution pattern. It is therefore important that if teachers are to correctly and consistently assess student work, they need to be able to detect patterns such as the one that can be found in Chris's work.

A second implication for teacher education is that the CCSS mathematical practice concerning viable arguments needs to receive detailed attention; not just as a stand-alone mathematical practice, but also in relation to the other mathematical practices. For instance, another mathematical practice recommends that students should "recognize patterns and make use of structure." In this study none of the participants recognized a possible pattern or a structure in Chris's argument consisting of the three first examples for the conjecture in the first evaluation item. The pattern could easily be extended into a general argument, but the participants failed to notice this. A stronger understanding of patterns as a foundation for viable arguments might have changed their perspective. Attention needs to be paid to how the mathematical practices interact in the subject of mathematics, so that they are not just treated as individual practices.

Additionally, this study revealed that viable argument is a multi-faceted term; the CCSS description contains a lot of information and mentions several different possible argument representations. Pre-service teachers need to be exposed to examples of all of these types of arguments. From Ryan's visual argument in the first evaluation item, it was evident that a visual representation of a mathematical result was unfamiliar to the participants and they all commented on the visual nature or the creativity of this argument. This type of argumentation is common in the education of K-8 teachers, and these types of arguments should be presented to secondary teachers as well.

Recommendations for Future Research

This study addressed understanding of viable arguments and argument construction and evaluation among pre-service secondary teachers with a strong background in proof, reasoning, and argumentation.

The Influence of Coursework

The results of this study show that teachers with the common background described above can successfully apply their educational experiences with proof to the analysis of argumentation structure when they evaluate the work of students. This does not mean that other groups will do the same, leading to an opportunity to examine whether different combinations of mathematical coursework and content knowledge support these practices. A related question is whether extended experience with formal proof in advanced mathematics courses creates a bias toward approaching problems from

a formal proof perspective rather than the more open-minded view implied by constructing viable arguments.

Also, since the participants were pre-service teachers they did not have a lot of experience with grading and evaluating authentic student work. Future research can contrast these findings with how in-service teachers, who have substantial experience evaluating student work, interpret, construct, and evaluate viable arguments. More study is also needed on how pre-service elementary teachers, who do not have the same background in mathematical proof, reasoning, and argumentation, construct and evaluate viable arguments, and what kinds of tasks are appropriate at lower grade levels.

The participants in this study expressed views that examples and empirical evidence are first and foremost used to gain understanding of a problem and to search for a counterexample. However, examples also have great potential power in explaining results and uncovering mathematical structures. More research is needed on the roles of examples in the proving process to discover if pre-service teachers recognize the importance of examples in mathematics, not only in proof production but also in the teaching of mathematics.

The participants' ability to construct arguments was clearly affected by their familiarity with the subject matter, but also by their mathematical confidence. It is important to keep in mind that the participants in this study were college students who have entered into a teaching preparation program specifically to become mathematics teachers. Even so, their confidence affected how they constructed and evaluated arguments. For instance, David's lack of confidence was apparent throughout the

interview, and he failed to provide viable arguments for two of the items. Similarly, Bridget's approach to the calculus item was hesitant, and she failed to recognize that she had produced an argument close to a valid proof, instead continuing on with a different approach. A well-designed research study might help teachers and teacher educators understand how confidence affects the argumentation process.

Concluding Statements

This study investigated pre-service teachers' understanding of the term "viable argument" as well as the strategies they employed to construct arguments and critique the reasoning of others. The study revealed that pre-service teachers with a strong background in proof understand viable argument as a mathematical practice where the main purpose is to explain one's reasoning. The study showed that the participants compared this to mathematical proof and that they recognized both similarities and differences between these two terms. The participants were found to have the understanding of mathematical proof, reasoning, and argumentation that will be needed to implement the recommendations of the Common Core Standards. Their goal and challenge as classroom teachers will be to convey this same understanding to their students.

REFERENCES CITED

Auerbach, C. F., & Silverstein, L. B. (2003). Qualitative data: An introduction to coding and analysis. New York: New York University Press.

Balacheff, N (1988). Aspects of Proof in pupil's practice of school mathematics. In Pimm, D. (Ed). Mathematics, Teachers, and Children. Hodder & Stoughton

Ball, D.L., Hoyles, C., Jahnke, N.H., & Movshovitz-Hadar, N. (2002). The Teaching of Proof. In L. I. Tatsien (Ed.), *Proceedings of the International Congress of Matematicians,* Vol. III, p. 907–920. Beijing: Higher Education Press.

Blum, W., & Kirsch, A. (1991). Preformal Proving: Examples and Reflections. *Educational Studies in Mathematics.* 22(2), 183-203 Auerbach, C. F., & Silverstein, L. B. (2003). Qualitative data: An introduction to coding and analysis. New York: New York University Press.

Balacheff, N (1988). Aspects of Proof in pupil's practice of school mathematics. In Pimm, D. (Ed). Mathematics, Teachers, and Children. Hodder & Stoughton

Ball, D.L., Hoyles, C., Jahnke, N.H., & Movshovitz-Hadar, N. (2002). The Teaching of Proof. In L. I. Tatsien (Ed.), *Proceedings of the International Congress of Matematicians,* Vol. III, p. 907–920. Beijing: Higher Education Press.

Blum, W., & Kirsch, A. (1991). Preformal Proving: Examples and Reflections. *Educational Studies in Mathematics.* 22(2), 183-203

Clement, John J. (2000)."Analysis of clinical interviews: Foundations and model viability", In Lesh, R. and Kelly, A. (Eds) ***Handbook of research methodologies for science and mathematics education.*** Hillsdale, NJ: Lawrence Erlbaum.

Common Core State Standards (2011). Available online at http://www.corestandards.org/assets/CCSSI_Math%20Standards.pdf

Conference Board of the Mathematical Sciences (2001) The Mathematical Education of`Teachers. AMS MAA

Conference Board of the Mathematical Sciences (2012). The Mathematical Education of`Teachers II. AMS MAA

Corbin, J.,& Strauss, A. (2008). *Basics of Qualitative Research (3rd ed),* London: Sage

de Villiers, M.D. (1999). Rethinking Proof with the Geometer's Sketchpad. Emeryville, CA: Key Curriculum Press.

Durand-Gurrier, V., Boero, P., Douek, N., Epp, S.S., Tanguay, D. (2012). Argumentation and Proof in the Mathematics Classroom. In Hanna, G., & deVilliers, M. (Eds). *Proof and Proving in Mathematics Education.* (pp. 349-367): Springer

Epp, S.S. (2003). The Role of Logic in Teaching Proof. *American Mathematical Monthly, 110*(10), 886-889.

Erlwanger, S.H. (1973). Benny's Conception of Rules and Answers in IPI Mathematics. *Journal of Children's Mathematical Behavior,* 1(2), 7-26. In Carpenter, T.P., Dossey, J.A. & Koehler, J.L. (Ed.s) *Classics in Mathematics Education Reasearch* (2004).Reston, VA: NCTM

Fischbein, E. (1982). Intuition and Proof, *For the Learning of Mathematics* 3(2), 9-18,24

Golafshani, N. (2003). Understanding Reliability and Validity in Qualitative Research. *The Qualitative Report,* 8(4), 597-607

Goldin, G.A. (2000). A Scientific Perspective on Structured, Task-Based Interviews in Mathematics Education Research. In Lesh, R. and Kelly, A. (Eds) **Handbook of research methodologies for science and mathematics education.** Hillsdale, NJ: Lawrence Erlbaum.

Hanna, G. (1998). More than Formal Proof, *For the Learning of Mathematics*, 9(1), 20-23

Hanna, G. (2000a). Proof, Explanation and Exploration: An Overview. *Educational Studies in Mathematics, 44*, 5-23

Hanna,G. (2000b). A Critical Examination of Three Factors in the Decline of Proof, *Interchange,* 31(1), 21-33

Harel, G., & Sowder, L. (2007). Toward Comprehensive Perspectives on the Learning and Teaching of Proof. In F.K. Lester, Jr (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 805-842). Charlotte, N.C.: Information Age Publishing.

Healy, L., & Hoyles, C. (2000). A Study of Proof Conceptions in Algebra, *Journal for Research in Mathematics Education* 31(4), 396-428

Hill, H.C, Ball, D.L., & Schilling, S.G. (2008). Unpacking Pedagogical Content Knowledge: Conceptualizing and Measuring Teachers' Topic-Specific Knowledge of Students. *Journal for Research in Mathematics Education,* 39(4), 372-400

Hoepfl, M.C. (1997). Choosing Qualitative Research: A Primer for Technology Education Researchers. *Journal of Technology Education,* (9)1. http://scholar.lib.vt.edu/ejournals/JTE/v9n1/hoepfl.html

Hoyles, C. (1997). The curricular shaping of students' approaches to proof. *For the Learning of Mathematics, 17(1),* 7-16

Jacobs, J.K., Hiebert, J., Bogard Givvin, K., Hollingsworth, H., Garnier, H., & Wearne, D. (2006). Does Eigth-Grade Mathematics Teaching in the United States Align with the NCTM Standards? Results from the TIMMS 1995 and 1999 Video Studies. *Journal for Research in Mathematics Education,* 37(1), 5-32

Knuth, E.J.(2002a). Secondary school mathematics teachers' conception of proof. *Journal for Research in Mathematics Education, 33*(5), 379-405.

Knuth, E.J. (2002b). Teachers' conception of Proof in the Context of Secondary School Mathematics. *Journal of Mathematics Teacher Education, 5,* 61-82.

Lincoln, Y. S., & Guba, E. G. (1985). Naturalistic inquiry. Beverly Hills, CA: Sage

Nardi, E., Biza, I., & Zachariades, T. (2012). Warrants Revisited: Integrating Mathematics Teachers' Pedagogical and Epistemological Consideration into Toulmin's Model for Argumentation. *Educational Studies in Mathematics,* 79(2), 157-173

National Council of Teachers of Mathematics [NCTM]. (2000). *Principles and Standards for School Mathematics.* Reston, VA: National Council of Teachers of Mathematics

Martin, W.G., & Harel, G. (1989). Proof Frames of Preservice Elementary Teachers. *Journal for Research in Mathematics Education, 20 (1),* 41-51.

Martinez, M. V., Brizuela, B. M., & Castro Superfine, A. (2011). Integrating Algebra and Proof in High School Mathematics: An Exploratory Study. *Journal of Mathematical Behavior*, 30, 30–47.

Mason, J., Burton, L., & Stacey, K. (1981). *Thinking mathematically*. London, UK: Addison Wesley

Morris, A.K., (2002). Mathematical Reasoning: Adults' Ability to Make the Inductive Deductive Distinction. *Cognition and Instruction, 20(1),* 79-118

Morris, A.K., (2007). Factors Affecting Pre-Service Teachers' Evaluation of the Validity of Students' Mathematical Arguments in Classroom Contexts. *Cognition and Instruction, 25(4),* 479-522

Patton, M.Q. (2002). *Qualitative Evaluation and Research Methods (3rd Ed).* Thousand Oaks, CA: Sage

President's Council of Advisors on Science and Technology (2011) http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-stem-ed-final.pdf

President's Council of Advisors on Science and Technology (2012) http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-engage-to-excel-final_2-25-12.pdf

Raman, M. (2001). Beliefs about Proof in Collegiate Calculus. In Proceedings of the Annual Meeting of the North American Chapter of the International Group for the Psychology for Mathematics Education (pp. 595-603)

Raman, M (2003). Key Ideas: What Are They and How Can They Help Us Understand How People View Proof? *Educational Studies in Mathematics, 52,* 319-325.

Reid, D. A. & Knipping, C. (2010). Proof in Mathematics Education, Research, Learning and Teaching, Rotterdam, The Netherlands: Sense Publishers.

Riley, K.J. (2003). An investigation of prospective secondary mathematics teachers' conception of proof and refutations. Doctoral dissertation. Montana State University.

Rowland, (2001). Generic proofs in number theory. In Campbell, S., and Zazkis, R. (Eds.). *Learning and teaching number theory: Research in cognition and instruction*. (pp. 157-184). Westport, CT: Ablex Publishing.

Rowland, T. (2008). The purpose, design and use of examples in the teaching of elementary mathematics. *Educational Studies in Mathematics, 69(2), 149-163*

Shaughnessy, J.M. (2004). Perspective on "Benny's Conception of Rules and Answers in IPI Mathematics". In In Carpenter, T.P., Dossey, J.A. & Koehler, J.L. (Ed.s) *Classics in Mathematics Education Reasearch* (2004). Reston, VA: NCTM

Schifter, D. (2009). Representation-based Proof in the Elementary Grades. In Stylianou, D.A., Blanton, M.L., Knuth, E.J. (Eds.). (2009). Teaching and Learning Proofs Across the Grades: A K-16 Perspective. (pp.71-86). New York, N.Y.: Routledge

Selden, A., & Selden, J. (2003). Validations of Proofs Considered as Texts: Can Undergraduates Tell Whether an Argument Proves a Theorem. *Journal for Research in MathematicsEducation.* 34(1), 4-36

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher,* 15(2), 4–14.

Sowder, L. and Harel, G. (1998). Types of Students' Justifications, *The Mathematics Teacher* 91(8), 670-675

Staples, M.E., Bartlo, J., & Thanheiser, E. (2012). Justification as a teaching and learning: Its (potential) multifaceted role in middle grades mathematics classrooms. *The Journal of Mathematical Behavior,* 31(4), 447-462

Stigler, J.W.,& Hiebert, J. (1999). *The Teaching Gap: Best Ideas from the World's Teachers for Improving Education in the Classroom.* New York, N.Y.: The Free Press

Stylianides, A.J.(2007a). Proof and Proving in School Mathematics. *Journal for Research in Mathematics Education, 38,* 289-321

Stylianides, A.J.(2007b) The Notion of Proof in the Context of Elementary School Mathematics. *Educational Studies in Mathematics, 65,* 1-20

Stylianides, G.J. (2008). An Analytic Framework for Reasoning-and-Proving. *For the Learning of Mathematics, 28,* 9-16

Stylianides, A.J.,& Ball, D.L. (2008). Understanding and describing mathematical knowledge for teaching: knowledge about proof for engaging students in the activity of proving. *Journal of Mathematics Teacher Education,* 11(4), 307-332

Stylianides, A.J.,& Stylianides, G.J. (2009a). Proof Constructions and Evaluations. *Educational Studies in Mathematics, 72*, 237-253.

Stylianides, G.J., & Stylianides, A.J. (2009b). Facilitating the Transition from Empirical Arguments to Proof. *Journal for Research in Mathematics Education,* 40(3), 314 352

Stylianides, G.J. (2009). Reasoning-and-Proving in School Mathematics Textbooks. *Mathematical Thinking and Learning,* 11(4), 258-288

Stylianou, D.A., Blanton, M.L., Knuth, E.J. (Eds.). (2009). Teaching and Learning Proofs Across the Grades: A K-16 Perspective. New York, N.Y.: Routledge

Tall, D., & Vinner, S. (1981). Concept Image and Concept Definition in Mathematics with Particular Reference to Limits and Continuity. In Carpenter, T.P., Dossey, J.A. & Koehler, J.L. (Ed.s) *Classics in Mathematics Education Reasearch* (2004). Reston, VA: NCTM

Varghese, T. (2009). Secondary-level Student Teachers' Conceptions of Mathematical Proof. *IUMPT:The Journal,* (1).http://www.k12prep.math.ttu.edu/journal/contentknowledge/varghese01/article.pdf

Varghese, T. (2011). Considerations Concerning Balacheff's 1988 Taxonomy of Mathematical Proofs. *Eurasia Journal of Mathematics, Science & Technology Education,* 7(3), 181-192

Walter, J.G., & Barros, T. (2011). Students build mathematical theory: semantic warrants in argumentation. *Educational Studies in Mathematics,* 78 (3), 323-342

Waring, S. (2000). Can you prove it? Developing Concepts of Proof in Primary and Secondary School, Mathematical Association

Weber, and Alcock (2004). Semantic and Syntactic Proof Production. *Education Studies in Mathematics,* 56 (2), 209-234

Weiss, M., Herbst, P., and Chen, C. (2009). Teachers' Perspectives on "Authentic Mathematics" and the Two-Column Proof Form. *Educational Studies in Mathematics, 70,* 275-293

Williams, S. R. (1991). Models of Limit Held by College Calculus Students. *Journal for Research in Mathematics Education,* 22(3), 219-236

Yackel, E. (2002). What we can learn from analyzing the teacher's role in collective argumentation. *Journal of Mathematical Behavior,* 21(4), 423-440

Yopp, D.A. (2009). From Inductive Reasoning to Proof. *Mathematics Teaching in the Middle School,* 15(4), 286-291

Yopp, D.A. (2010). Developing an Understanding of Logical Necessity. *Teaching Children Mathematics,* 16(7), 410-422

Yopp, D.A. (2011). How some Research Mathematicians And Statisticians use Proof in under-graduate Mathematics. *Journal of Mathematical Behaviour,* 30(2), 115-130

Yopp, D.A, Burroughs E.A., Lindaman B.J (2011) Why it is important for in-service elementary teachers to understand the equality .999…=1. *The Journal of Mathematical Behavior* 30(4), 304-318

APPENDICES

APPENDIX A

INTERVIEW PROTOCOL

Interview Protocol: Session 1 – Student _____

Hello, my name is Kim Nordby. I am a Mathematics Education doctoral student here at MSU. I'm glad you have agreed to participate in this interview, and I hope you'll find it interesting.

Over the next few days I'll be interviewing several of the students enrolled in M329 this semester to find out what ideas you have about mathematical argumentation. Your answers will help the faculty make decisions about our mathematics education program so that we can better serve future students' needs.

During the interview we'll look at several mathematics problems, some that you will work on yourself and some where you will examine other students' work. There is no right or wrong response to the questions I ask about your perceptions and mathematical thinking. My purpose is to learn more about how a pre-service teacher in your position views and interprets argumentation in mathematics.

If any question makes you uncomfortable, you do not have to answer. You can also end the interview at any point and for any reason. So, if you change your mind and want to stop just let me know, and we'll stop right then.

With your permission, I would like to record our conversation so I can watch it later and reflect more carefully on your comments. I want to assure you that your answers will be kept as private and confidential as possible. The recording of this interview will be kept safe and will viewed only by me and possibly my advisor, and it will be erased after the study is complete. When I write up my results, responses from participating students will be made anonymous so that no one can identify the person giving a specific answer. If I quote or refer to your responses in my report, your name will not be mentioned and no one will be able to guess your identity based on your answers. Do I have your permission to record this session?

Again, I want to thank you and emphasize that your participation is very important to me both for my own research and to help us guide future improvements to MSU teacher preparation.

I’ve prepared a series of questions built around a set of mathematical tasks. I am following an outline, but may sometimes ask additional questions to clarify or expand on your responses.

TASK A:
Let's begin with a geometry task. While considering this problem, imagine that you have access to whatever resources, tools, manipulatives, or software you might find useful.

Question1: Describe how you would construct an argument to support the following <u>true</u> statement:

The sum of the angles in a triangle is 180°.

Potential prompts:

- *Why did you choose that particular approach?*
- *Are there other methods you might have used?*
- *Would you use the same approach to convince a 5^{th} grader? A mathematician?*

TASK B:
Are you familiar with the Common Core State Standards for Mathematics?

Potential prompt:

- *The Common Core State Standards have been adopted by most states as guidelines for K-12 mathematics curriculum and instruction. Most likely, the mathematics you teach in your future career will be governed by the Common Core.*

The Common Core includes eight Standards for Mathematical Practice. One of these states that students should be able to "construct viable arguments."

Question2: What do you think is meant by the term "viable argument"?

Potential prompts:

- *What is a viable argument in support of a mathematical conjecture? Over an infinine set?*
- *What does it mean to provide a viable argument?*
- *Definitions of "viable" include: practical; workable; capable of success or continuing effectiveness;* capable of working, functioning, or developing adequately

Question3: What characteristics qualify an argument as viable?

Potential prompt:

- *What features do you look for in an argument relating to a mathematical conjecture to determine if it is viable or not?*

Question4: What is the difference between a proof and a viable argument?

Question5: What is the role of "viable arguments" in school mathematics?

Question 6: When should students encounter viable arguments?

- What is a viable argument in elementary school?
- What is a viable argument in secondary school?

TASK C:

The Common Core offers more detail about what it means to construct viable arguments. I'd like you to read this description while reflecting on the answer you gave a moment ago.

(Hand the student a copy)

Question7: How do you interpret this description of "viable argument" provided by the writers of the CCSS?

Question8: Does this description change any of your previous thoughts of what is meant by a "viable argument"?

Potential prompt:

- *Does anything in this description surprise or particularly interest you?*

3. Construct viable arguments and critique the reasoning of others.

Mathematically proficient students understand and use stated assumptions, definitions, and previously established results in constructing arguments. They make conjectures and build a logical progression of statements to explore the truth of their conjectures. They are able to analyze situations by breaking them into cases, and can recognize and use counterexamples. They justify their conclusions, communicate them to others, and respond to the arguments of others. They reason inductively about data, making plausible arguments that take into account the context from which the data arose. Mathematically proficient students are also able to compare the effectiveness of two plausible arguments, distinguish correct logic or reasoning from that which is flawed, and—if there is a flaw in an argument—explain what it is. Elementary students can construct arguments using concrete referents such as objects, drawings, diagrams, and actions. Such arguments can make sense and be correct, even though they are not generalized or made formal until later grades. Later, students learn to determine domains to which an argument applies. Students at all grades can listen or read the arguments of others, decide whether they make sense, and ask useful questions to clarify or improve the arguments.

TASK D:

Let's return to the problem and angle sums in a triangle.

Question9: In your opinion, is the argument you constructed a "viable argument?"

Potential follow-up: Are there any situations where your argument would not be viable?

Question10: What features need to be present for an argument to be considered a "viable argument"?

TASK E:
Let's move on to another mathematics problem, this time from number theory.

Question 11: Please demonstrate how you would construct a viable argument to prove or disprove the following statement:

The sum of two consecutive odd numbers is divisible by four.

Question 12: Now that you've spent some time on this problem, tell me about your thought process while working on your argument.

- How did you go about investigating the conjecture?
- What choices or decisions did you make along the way?
- What steps did you take, and in what order, to construct a viable argument?
- What features of your work allow you to classify your argument as a "viable argument"?

TASK F:
I'm now going to ask you to examine some student work on the following conjecture:

The sum of two consecutive odd numbers is divisible by four.

In no particular order, I am giving you five arguments written by students supporting the truth of this conjecture. Please take a moment to look them over and consider their "viability" before we continue with the interview. *(Pause…)*

Recall that the Common Core suggest students should be able to "critique the reasoning of others." I'd like you to share your critique of these student arguments. You may consider all five responses together, or refer to specific responses.

Question13: Among all five responses, which do you consider to represent viable arguments? Why?

Question14: What are the merits of the arguments you consider viable?

Question15: What are the shortcomings of the arguments you do not find viable?

Question 16: For the arguments that you said are not viable, what must be added to the arguments to make them viable?

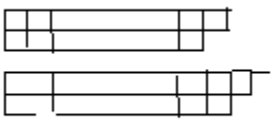
Answer 1: 1+3=4, 5+7=12
7+9=16 11+13=24
53+55=108 123+125=248
We see that all the answers are divisible by four so the sum of two consecutive odd numbers is divisible by four

Answer 2: 1+3=4
3+5=8
5+7=12
We see that the sum of two consecutive numbers is divisible by four.

Answer 3: Consider two consecutive odd numbers. One of these numbers must be one more than a multiple of four and the other must be one less than a multiple of four. For example, 25 is one more than a multiple of four and 27 is one less than a multiple of four. So, when you add them together, you get a multiple of four. 25+27=52, which is a multiple of four.

Answer 4: If you add two odd numbers the answer is always even. When you add two even numbers, the answer is always a multiple of four. So, when you add two consecutive odd numbers the answer is always a multiple of four.

Answer 5:



When these two consecutive odd numbers are added together, we see that the one box left over in the larger odd number will fill in the missing box in the smaller odd number creating a new box with four rows and no boxes left over. Since there are four rows, the number is a multiple of four. This means that the sum of two consecutive odd numbers must be a multiple of four

Interview Protocol: Session 2 – Student _____

Once again, I would like your permission to record our conversation so I can watch it later and reflect more carefully on your comments. I want to assure you that your answers will be kept as private and confidential as possible. The recording of this interview will be kept safe and will viewed only by me and possibly my advisor, and it will be erased after the study is complete. When I write up my results, responses from participating students will be made anonymous so that no one can identify the person giving a specific answer. If I quote or refer to your responses in my report, your name will not be mentioned and no one will be able to guess your identity based on your answers. Do I have your permission to record this session?

TASK G:
This task provides you with another opportunity to work on a conjecture yourself. I invite you to talk aloud as you work - I'm interested in understanding your thought process as you approach the task.

Question17: Provide a viable argument for or against the following conjecture:

If a divides b and b divides c, then a divides c.

Question18: Let's talk about how you went about constructing your argument.
If these features weren't transparent while the student worked, ask:

- How did you go about investigating the conjecture?
- What choices or decisions did you make along the way?
- What steps did you take, and in what order, to construct a viable argument?
- What features of your work allow you to classify your argument as a "viable argument"?

Question19: Are there any contexts where your argument would not qualify as "viable"?

Potential prompts:

- *Would this argument be considered viable in a 5^{th} grade classroom?*
- *Would this argument be considered viable in a graduate mathematics class?*

Question 17b (back-up, if needed): Provide a viable argument for or against the following conjecture:

The derivative of an even function is odd

(if used followed by questions 18 & 19)

TASK H:
In this final task, you'll again examine student work. The rule for testing a whole number for divisibility by 3 is stated as follows:

A number is divisible by 3 if the sum of the digits is divisible by 3.

In no particular order, I am giving you four arguments written by students supporting the truth of this conjecture. Please take time to analyze each argument and consider its "viability" before we continue with the interview. *(Pause…)*

As before, I'd like you to critique of these students' reasoning. Please proceed in any way that makes sense to you. *(Allow interviewees to express themselves before offering the questions below.)*

Question20: Among all five responses, which do you consider to represent viable arguments? Why?

Question21: What are the merits of the arguments you consider viable?

Question22: What are the shortcomings of the arguments you do not find viable?

Question23: For the arguments that you said are not viable, what must be added to the arguments to make them viable?

Argument 1:

Number	Sum of Digits	Number divisible by 3?	Sum of digits divisible by 3?
12	3	Yes	Yes
13	4	No	No
360	9	Yes	Yes
361	10	No	No
137,541	21	Yes	Yes
157,541	23	No	No

We see that the rule is true.

Argument 2:
Let a be any three digit number, with digits x,y, and z. By the place-value concept, a=100x+10y+z. This equality can be writte a=(99x+x)+(9y+y)+z. By the commutative and associative properties we get a=(99x+9y)+(x+y+z). Notice that the expression 99x+9y is always divisible by 9, and therefore also by 3. Now, if the second expression, which is the sum of the number's digits, is divisible by 3, then we get that the number itself is divisible by 3.

Argument 3:
Let a be any whole number such that the sum of its digits is divisible by 3. Assuming its digits are x,y, and z, then a=xyz. Since x+y+z is divisible by three, also xyz is divisible by three. Therefore a is divisible by 3

Argument 4:
Consider 756. This number can be represented as follows: 756= 7x100+5x10+6x1. This can be rewritten 756= (7x99+7)+ (5x9+5)+6. By the commutative and associative properties, we get 756=(7x99+5x9)+(7+5+6). Notice that the expression 7x99+5x9 is always divisible by 9, and therefore also by 3. Now if the second expression which is the sum the number's digits, is divisible by 3, then we get that the number itself is divisible by 3.

Question 24: Please tell me about your mathematical background. Which college level mathematics classes have you completed?

Question 25: Who were the instuctors in M242 and M328?

Question 26: Is there anything you would like to ask me, or to add regarding your views on constructing viable arguments and critiquing the reasoning of others?

This concludes our interview. Thank you again for your time and willingness to participate!

APPENDIX B

CONSENT FORM WITH IRB APPROVAL

Videorecording Consent Form

Dear MSU Student: March 2013

We are asking your consent to participate in a videorecorded interview about the mathematics you are learning in your teacher preparation program at MSU. The recording is for Kim Nordby's doctoral dissertation in conjunction with research on the mathematics education programs offered at MSU. This research will help us better understand the best teaching practices related to examples, arguments, and proof and possibly identify areas of improvement. The recordings will be used primarily for data analysis. Data collected from these interviews may also be used in research journals and other publications. All of this data will be anonymously reported - your name will not be associated with data recorded from the interviews. If any question makes you uncomfortable, you do not have to answer. You can also end the interview at any point and for any reason.

Video excerpts may be included in research presentations at conferences and workshops. You can agree or decline to allow your video images to be used in this way.

There will be two videorecorded sessions, the first lasting less than 45 minutes, the second less than 30 minutes. There is no risk to you in this process, and gathering such information is considered a normal part of teachers' work in the classroom. Your grades will not be affected in any way by your participation or your refusal to participate in this project, and the information you provide is in no way connected to course assessment. The recordings are for the purpose of dissertation research conducted by Kim Nordby and will be viewed only by him, his faculty advisor, and a second researcher.

Your permission for us to interview and record you mathematical thinking in this research is voluntary. There will be no loss of teaching or learning opportunity if you decide not to participate.

Please sign and date the consent form indicating whether you agree to participate or not. Please keep a copy for your records and return one copy to Kim Nordby. If you have questions about the project, please contact nordby@math.montana.edu or (406) 595-8007.

Sincerely,

Kim Nordby

_____

Print Name

By signing here, you agree to participate in the interview, have your interview videorecorded, and allow data from the interview to be used anonymously in research and publication.

_____ _____

Signature Date

By signing below, you further agree to allow video excerpts of your interview to be potentially included in research presentations and workshops.

_____ _____

Signature Date

APPROVED
MSU IRB
02/28/2013
Date Approved