THE USE OF EXAMPLES IN A TRANSITION-TO-PROOF COURSE

by

Sarah Elizabeth Mall Hanusch, M.A., B.S.

A dissertation presented to the Graduate Council of
Texas State University in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
with a Major in Mathematics Education
August 2015

Committee Members:

Samuel Obara, Chair

Alexander White

Sharon Strickland

Keith Weber

ACKNOWLEDGMENTS

First, I want to thank the members of my committee, for their support through my entire time in the program. I would not have finished this program without their encouragement and insight.

Second, I want to thank my family and friends. My parents raised me to value education, and provided me with the resources to be successful. I owe everything to them. My friends have cheered and cried with me throughout this journey, and I am thankful for their companionship. I am especially grateful to Yuliya for being my writing partner over the last year; she always provided a willing ear and quality feedback.

Finally, I need to thank my husband, Chris. I would not have started this journey without his encouragement and approval. He knew that I needed to prove to myself that I am capable of earning a doctorate, and has willingly put the rest of our life on hold so I could. I am so thankful to have him as a partner in life. I love you, Chris.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

This study investigates the ways that undergraduate students use examples in their transition-to-proof course, and the influence that the instructor had on the students' decisions to use examples. Observations and interviews were conducted with the instructor and a sample of students to investigate the connections between the teaching and learning of examples in this proof writing course.

CHAPTER 1

INTRODUCTION

Many undergraduates do not understand how to develop and write proofs (Moore, 1994; Weber, 2001). Learning to write proofs is a complicated process, and results in students developing a variety of beliefs about what constitutes a proof and how one goes about constructing a proof (Harel & Sowder, 2007). Using examples is one possible strategy in the proof writing process. Examples can be used for several purposes when developing and proving conjectures (Alcock & Inglis, 2008; Alcock & Weber, 2010; Lockwood, Ellis, & Knuth, 2013). As a consequence of their work, the researchers cited above theorize that if undergraduate mathematics students are introduced and instructed in generating and using examples in order to construct proofs, then the students' ability to construct proofs may improve.

The term example can have many different meanings in mathematics (Watson & Mason, 2002). When a student asks an instructor for an example, the student often wishes to see a demonstration of a technique for solving a problem, or a specific execution of an algorithm. In other circumstances, an example might refer to a list of types of problems with known solutions, a list of classes of objects, or objects satisfying a given definition. Within this study, the term example is limited to a mathematical object which satisfies specific characteristics and illustrates a definition, concept or statement (Moore, 1994). This definition excludes sample proofs, e.g., demonstrations of the direct proof technique or proving by induction. Alcock and Weber (2010) claim that this definition of example is "probably the most common intended meaning of the term when it is used by mathematicians and mathematics educators in the context of proof-oriented mathematics" (p. 2).

There is a long history of mathematics instructors advising their students to construct examples as a problem solving strategy (Sandefur, Mason, Stylianides, & Watson, 2012). For instance, Polya (1957) suggested that one strategy for solving a problem is to look at simple cases and extrapolate to general assertions.

One reason that students of mathematics and mathematicians use examples is to extend their conceptual knowledge of mathematical objects (Alcock, 2009; Alcock & Inglis, 2008; Watson & Mason, 2005). In other terms, thinking about examples can expand a learner's current concept image, as defined by Vinner (1991). The collection of examples known to a learner defines the learner's example space for the concept (Watson & Mason, 2005). Ideally, this example space will include many different examples of the concept. Michener (1978) argued that a full understanding of a mathematical topic includes knowledge of start-up examples, reference examples, model examples and counterexamples. Michener's scheme claims that start-up examples introduce a new subject or concept, reference examples are standard examples that are widely applicable to a subject, model examples are generic and represent important aspects of the concept, and counterexamples show a statement is not true.

One particular type of example that is useful for proof writing is a generic example (Alcock & Inglis, 2008; Harel & Tall, 1991; Mason & Pimm, 1984; Rowland, 2001). A generic example is a concrete example that is "presented in such a way as to bring out its intended role as the carrier of the general" (Mason & Pimm, 1984, p. 287). For instance, to prove that $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$, a prover could consider a specific case such as $n = 10$. By reversing the terms of the sum, adding each pair $1 + 10, 2 + 9, \ldots, 9 + 2, 10 + 1$, and then adding the pairs together and dividing by 2 the structure of the general argument is revealed through the specific. As such, this example satisfies the definition of a generic example.

Proof and reasoning are important aspects of mathematics and mathematical research, so it is important that undergraduate students in upper-division mathematics courses develop a good foundation in proving techniques. However, little is known about the role of instruction while students are learning how to write proofs. One goal of this work is to investigate the connection between students' proof writing processes and the instruction they received, especially with regards to the specific tool of using examples.

The course selected for this study is a semester long course in proof writing, titled Introduction to Advanced Mathematics. The purpose of this course is to introduce students to proof writing processes and strategies. The course is described as "[a]n introduction to the theory of sets, relations, functions, finite and infinite sets, and other selected topics. Algebraic structure and topological properties of Euclidean Space, and an introduction to metric spaces" (Texas State University-San Marcos, 2013).

Statement of the Problem

It is well documented that undergraduate students struggle when they start taking proof-based mathematics courses (Bills & Tall, 1998; Sowder & Harel, 2003; Weber, 2005b; Weber & Alcock, 2004). Weber (2001) classified the difficulties that undergraduates experience while learning to prove into three categories: an inadequate conception of what constitutes a mathematical proof, misunderstanding or misusing a definition or example during proof construction, and a lack of strategic knowledge. Weber (2001) defines strategic knowledge as "heuristic guidelines that they can use to recall actions that are likely to be useful or to choose which action to apply among several alternatives" (p. 111). Heuristics are difficult to teach, but students typically do not learn them unless an attempt is made to teach them (Lester, 1994). As such, there is a need to study the instruction of strategic knowledge and to analyze the effect, if any, the instruction has on the students' strategic knowledge. Questions about this topic were posed as directions for future research in Weber (2005a), including "[h]ow do the actions of the professor or the presentations in textbooks influence the ways in which students attempt to construct proofs?"(p.359). Answering this question is one of the goals of the present study.

Using examples while developing a proof is one particular strategy that a prover may utilize. This strategy is sometimes necessary, such as when a person wants to disprove a statement with a counterexample; sometimes optional, such as when an example generates the idea for a proof; and other times ineffective, such as when a prover can write an accurate proof immediately. Furthermore, the effectiveness of this strategy depends on which examples are chosen and the thinking processes of the prover. Within a proof construction, a prover should consider whether or not an example is useful.

Purpose of the Study

The purpose of this study is three-fold: 1) to develop a theory of effective examples use by undergraduate mathematics students during a transition-to-proof course, 2) to document how the instructor of a transition-to-proof course presented the strategies of example use to the students, and 3) to determine what impact, if any, that the instruction had on the students' effectiveness at using examples. Previous studies on undergraduate example usage on proof tasks have focused on the purposes for which students use examples and how students connect examples to their proof productions (Alcock & Weber, 2010; Lew, Mejia-Ramos, & Weber, 2013). This study will extend this work by looking at the entire process from when the students first indicates they will use an example until they complete their work on the problem.

Empirical studies on the teaching of upper-level undergraduate mathematics are scarce (Speer, Smith, & Horvath, 2010), however a few studies have begun to fill this gap (Fukawa-Connelly, 2012a, 2012b; Lew, Fukawa-Connelly, Mejia-Ramos, & Weber, 2015; Mills, 2014; Weber, 2004, 2012). Most of these studies focused on abstract algebra or real analysis courses and none occurred in transition-to-proof courses. In addition, the existing literature on the teaching of upper-level undergraduate mathematics courses have focused on either the content and effectiveness of lecture as an instructional strategy, or on the teaching of proof writing strategies as a whole. This study intends to limit its focus to one component (examples) and not just how the instructor uses examples during the lectures, but also what the instructor told the students about using examples themselves.

Finally, this study will compare the behavior of the students to the instruction provided. In most of the previous studies on instruction of upper-level undergraduate mathematics courses, the instructors were the focus and the students were not considered. By considering the instructor and her students simultaneously and asking the students about their perceptions of the course and the instructor about her perceptions of the students insight into the connections between the two can be established.

Significance of the Study

This study will contribute to the body of knowledge by exploring the connection between instruction and the students' decisions regarding example use during proof construction. A few studies have focused on the instruction that mathematics professors give to teach students proving techniques (Alcock, 2009; Mills, Brown, Larsen, Marrongelle, & Oehrtman, 2011). These studies used interviews with the instructors as the data collection technique. While interviews can provide valuable insight into the perceptions the professors have of their students, they provide only one view of classroom activities.

Conversely, some studies have looked at the ways undergraduate students use examples while constructing their proofs (Alcock & Weber, 2010; Iannone, Inglis, Mejia-Ramos, Simpson, & Weber, 2011). These studies used data from task-based interviews or an analysis of students' midterm exams supported by interviews. These studies focused on the approaches used by students, but again only used one view of classroom activities. The study proposed in this document will use in-class observations, supplemented by interviews, of both the instructor and the students to document both views of the classroom.

Definition of Terms

A list of terms used throughout this study is alphabetized below.

- Construction: When constructing an example, one often constructs an object and then verifies if it satisfies the properties to be an example. The term construction will apply to the object independent of the properties it holds. Once various properties are verified or refuted the construction can then be called an example, counterexample, non-example, boundary example, etc.

- Counterexample: A construction which satisfies the hypotheses of a statement, but not the conclusion is called a counterexample.

- Effective example use: A prover uses examples effectively when the example permits additional progress on the proof or results in a counterexample.

- Example: Any mathematical object satisfying a collection of characteristics that illustrates a concept, definition or statement (Moore, 1994). Once a construction has had all the requisite properties verified it can be called an example.

- Incomplete: A proof or example is classified as incomplete if a prover chose to end the process before completion. Typically this occurs if the prover deems themselves unable to continue at the present time or the session ran out of time.

- Incorrect: A proof or example is classified as incorrect if the prover believes their proof or example to be correct, but the researcher evaluates the item to be incorrect.

- Proof construction: Constructing a proof is "a mathematical task in which the prover is provided with some initial information (e.g. assumptions, axioms, definitions) and is asked to apply rules of inference (e.g., recall previously established facts, apply theorems) until a desired conclusion is deduced"(Weber, 2005a, p. 352). For this study, the proof construction will occur primarily in a written format, so the term proof writing may be used synonymously with proof construction in a written format. This definition does not preclude constructive or existence proofs, as these types of proof also use rules of inference to connect definitions to the statement that is being proved.

- Proof-related task: This term is used to describe the types of tasks asked of students in upper-division undergraduate courses. In this particular study these tasks include: proving statements, constructing examples, making conjectures, evaluating the truth of a statement, and validating the proofs of others.

- Prover: In this document, a prover is any individual working on a proof construction. However, the definition of prover used in this document is consistent with several other studies (Alcock, 2010; Fukawa-Connelly, 2012a;

Weber, 2001, 2002, 2005b, 2005a; Weber & Alcock, 2004).

- Strategic knowledge: The knowledge needed to find an appropriate and efficient strategy for solving a problem or developing a proof.

Theoretical Framework

This study will utilize grounded theory, a methodological technique developed by Glaser and Strauss (1967). Within this method, a researcher or a team collects and organizes data by constantly organizing the data into categories or themes (Charmaz, 2006). Merriam (2009) describes the constant comparative method as the process of "identifying segments in your data set that are responsive to your research questions"(p. 176), then "to compare one unit of information with the next" (p. 177), and eventually sorting all of the units into categories or themes. The process is "inductive and comparative" (Merriam, 2009, p. 175) because it occurs as the data is being collected and each new piece of data is compared and integrated into the categories as they are being developed. Once all of the data is collected the categories will be interpreted (Creswell, 2013). The theory developed accounts for the data collected, and is typically synthesized with relevant literature at the end of the process (Charmaz, 2006). The theory developed from the data in this study is introduced in the next section.

A theory of effective example usage in proof writing. Writing a proof can be a difficult process, in part because several different strategies could be implemented successfully to create a complete proof. A significant part of learning these strategies is evaluating which strategy to try, whether the strategy was implemented properly, and what conclusions can be drawn after implementation. Similar sentiments are established in the problem-solving literature; successfully solving problems depends on resource management (Carlson & Bloom, 2005; Polya, 1957; Schoenfeld, 1992). As Schoenfeld states "it's not just what you know; it's how, when, and whether you use it" (1992, p. 60). Constructing an example is one strategy that may be used during the process of proof writing. Undergraduate students, doctoral students and mathematicians use examples for several purposes including understanding a statement, evaluating the truth of a statement, and generating a proof (Alcock & Inglis, 2008; Alcock & Weber, 2010). Unfortunately, simply knowing that examples can be useful is insufficient for students to actually use examples effectively. Students use examples ineffectively by constructing examples that do not satisfy the given hypotheses, and the students also struggle with connecting the specific example to the formal language of proof (Alcock & Weber, 2010).

In order to use examples effectively during proof writing tasks, an individual needs knowledge in four phases: when to use an example, what purpose an example serves, how to construct the example, and what conclusions can be drawn from the example (see Figure 6). The first category, indicators of when to construct an

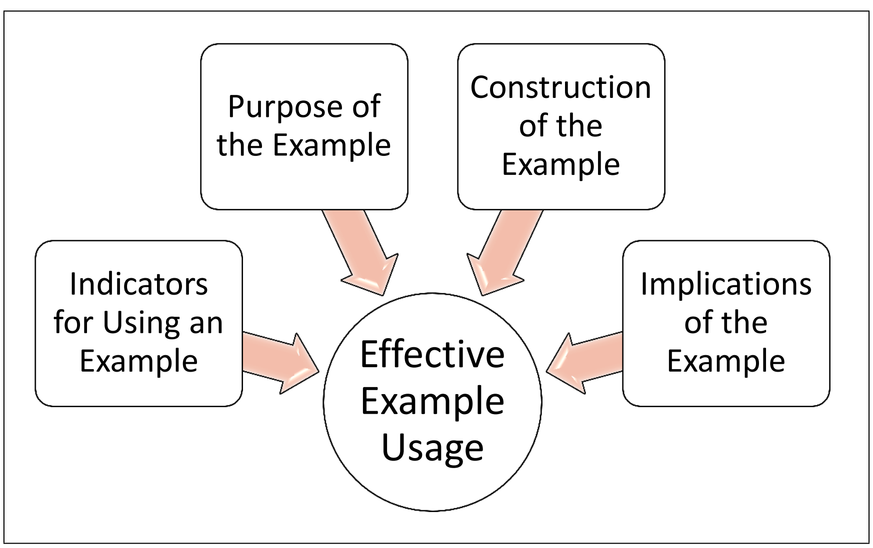

Figure 1: A model of effective example usage. The data in this study revealed four stages for effective example use: indicators, purposes, construction and implications.

example, includes characteristics of the problem and self-monitoring. Examples can provide insight at several stages in the proving process and are frequently used to form a conjecture and to verify the truth of an assertion (Lockwood, Ellis, Dogan, Williams, & Knuth, 2012). Consequently, some the phrasing of some problems (i.e. *prove or disprove*) may prompt provers to use examples when they otherwise may not have. Examples also provide provers the opportunity to make sense of mathematical definitions (Sandefur et al., 2012). When a prover is monitoring their own understanding of mathematical statements, the realization that a statement does not make sense may inspire the prover to construct an example to gain insight. In order to effectively use examples, a prover needs to know when an example might be useful.

Several purposes or uses of examples are explored in literature, including verifying the truthfulness of a claim, generating a counterexample, generalizing, making conjectures, understanding the statement of a claim or definition, and gaining insight into the argument for a proof (Alcock & Weber, 2010; Lockwood et al., 2012; Mills, 2014). The intended purpose of an example informs the construction of the example and the possible implications.

One way that students use examples ineffectively is by constructing objects that do not satisfy the given definitions or hypotheses (Alcock & Weber, 2010). These constructions are particularly troublesome when students classify these objects as examples or counterexamples and reach erroneous conclusions. Thus, to effectively use examples a prover needs to know how to accurately generate examples. Three strategies have been identified for generating examples: *trial and error*, *transformation* and *analysis* (Antonini, 2006; Iannone et al., 2011). *Trial and error* is the process of starting with recalled examples and testing the conditions for an example. The *transformation* strategy consists of starting with an example that satisfies some of the desired characteristics and adjusting the example until it satisfies all of them. The *analysis* strategy starts by assuming an example exists and continues with an analysis of the required properties. This analysis allows the prover to deduce additional properties until the desired example was recalled, the prover develops a procedure that constructs an example, or a contradiction arises. Antonini (2006) observed that graduate students in mathematics typically use these strategies in succession beginning with the *trial and error* strategy, then the *transformation* strategy, and only moving onto the *analysis* strategy when the other strategies are ineffective. A prover needs to know how to construct examples and how to verify that objects satisfy the required properties of an example. A prover may also find examples from *authoritarian* sources, such as textbooks, websites or other people. These examples are not technically constructions; however, a diligent prover will verify that examples from such sources truly satisfy all the desired conditions.

The final category for effective example usage is knowing what to do after using an example. A prover can make several different decisions after successfully constructing an example, including constructing another example or counterexample, reaching a conclusion regarding the truthfulness of a claim, using the example to write a formal proof of the claim, and disregarding the example altogether. Students often struggle with this component of using examples, which limits their effectiveness as a proof writing strategy (Alcock & Weber, 2010; Mejia-Ramos & Weber, 2015).

Research Questions

In this study, the following questions are addressed:

1. In what ways did the students use examples effectively and ineffectively on tasks during their transition-to-proof course? What aspects were implemented correctly and what aspect were implemented incorrectly?

2. How did the instructor teach students to effectively use examples?

    - How did the instructor model effective example use?
    - What did the instructor say (but not model) about using examples effectively?
    - How did the instructor design the course to teach this strategy?

3. What connections, if any, are found between the students' use of examples and the instruction given?

    - What behavior do the students attribute to the instruction?
    - What does the instructor expect the students to do with examples?
    - What alignment, if any, exists between the instructor's expectations of student behavior and the students' actual behavior?

CHAPTER 2

REVIEW OF LITERATURE

As described in the introduction, this study is focused on theorizing how students effectively use examples as a strategy during proof-related tasks, and how the instructor taught the students about the strategy. Due to the two components of this project, the review of literature over the next pages is divided into two large sections: literature concerning how individuals solve proof-related tasks and literature concerning how we teach students to solve proof-related tasks. Within each section, the relevant literature will be narrowed along a progression from general problem solving literature through proof writing literature and ending with examples literature.

Individuals Solving Proof-Related Tasks

Although a primary focus of transition-to-proof courses is writing proofs, some of the tasks assigned are not solely proofs. For instance, a task may ask for an example with certain properties. The amount of proof required depends how easy it is to verify the conditions hold. Other tasks include evaluating the truth of statements or validating the proofs of others. In the sections below, the literature on solving these various types of tasks for different populations is reviewed. Since all of these tasks are considered problems, the review begins with some of the literature on problem solving.

Problem solving literature. Problem solving is a focus of mathematics education (National Council of Teachers of Mathematics [NCTM], 2000) and as a consequence has been written about extensively (e.g. Carlson & Bloom, 2005;

Lester, 1994; Polya, 1957; Schoenfeld, 1992). Early studies focused on the features of tasks that led to difficulties, however now "there is a general agreement that problem difficulty is not so much a function of various task variables as it is the characteristics of the problem solver" (Lester, 1994, p. 664).

The process of solving problems has been studied repeatedly. Many studies on this topic trace their work back to Polya's (1957) influential work, where he describe the four phases of problem solving as 1) understanding the problem, 2) devising a plan, 3) carrying out the plan, and 4) looking back. He described this process as progressing linearly from one phase to the next. Another framework focuses on levels of behavior including a) *resources* the facts and procedures available to an individual, b) *control* how the individual chooses which resources to implement, and c) *belief systems* which are the attitudes an individual has about the nature of the discipline and the problems at hand (Schoenfeld, 1983). In a more recent framework, Carlson and Bloom (2005) have integrated these two frameworks with a process of orienting, planning, executing and checking, where the final three steps are repeated as needed, where resources, heuristics, affect and monitoring are utilized within each step. The types of resources, heuristics, affect and monitoring conducted depend on the step of the process (Carlson & Bloom, 2005).

Heuristics may be applied at any stage in the problem solving process, although a particular heuristic may only apply to one stage. A few common heuristics are found in Figure 2. Implementing these heuristics involves more than just being able to cite the rule; a problem solvers needs to evaluate which techniques might apply and then what might the implementation actually look like (Schoenfeld, 1980).

Another study on problem solving investigated the heuristic of determining key terms when solving problems. Lester and Garofalo (1982) found that many

Analysis (Understanding the problem)

1. Draw a diagram if at all possible.
2. Examine special cases, including: choosing special values, examining limiting cases, and setting integer parameters to 1,2,3,..., in sequence and looking for an inductive pattern.
3. Try to simplify the problem by exploiting "without loss of generality" arguments.

Exploration (Devising a plan)

1. Consider essentially equivalent problems:
   (a) Replacing conditions by equivalent ones.
   (b) Re-combining the elements of the problem in different ways.
   (c) Re-formulate the problem by
      i. changing perspective or notation
      ii. considering argument by contradiction or contrapositive
      iii. assuming you have a solution and determining its properties
2. Consider slightly modified problems:
   (a) Choose subgoals (obtain partial fulfillment of the conditions)
   (b) Relax a condition and they try to re-impose it
   (c) Decompose the domain of the problem and work on it case by case.
3. Consider broadly modified problems:
   (a) Construct an analogous problem with fewer variables.
   (b) Hold all but one variable fixed to determine that variable's impact.
   (c) Try to exploit related problems with similar forms, givens or conclusions.
   (d) Try to exploit both the results and the method of the easier problem.

Verifying Your Solution (Looking back)

1. Does your solution use all the pertinent data?
2. Can it be obtained differently?
3. Can it be used to generate something you know?

Figure 2: A list of some common heuristics for problem solving (Schoenfeld, 1980). This list has been edited for space and content considerations.

third and fifth graders believe that many problems can be solved by identifying *key words* that quickly indicates the correct strategy. This strategy can be effective for certain problems, but limits the types of problems an individual can solve.

Metacognition or self-monitoring. One definition of metacognition is

> "Metacognition" refers to one's knowledge concerning one's own cognitive processes and products or anything related to them, e.g., the learning-relevant properties of information or data.... Metacognition refers, among other things, to the active monitoring and consequent regulation and orchestration of these processes in relation to the cognitive objects on which they bear, usually in the service of some concrete goal or objective (Flavell, 1976, p. 232).

This definition has two components, the knowledge of cognition and the regulation of cognition. The knowledge of cognition concerns the processes, resources and strategies known to person, whereas, the regulation of cognition concerns monitoring one's understanding of a task and decision making during a problem solving process (Garofalo & Lester, 1985).

Schoenfeld discussed managerial problem-solving behaviors which include

> selecting perspectives and frameworks for a problem; deciding at branch points which direction a solution should take; deciding whether, in the light of new information, a path already taken should be abandoned; deciding what (if anything) should be salvaged from attempts that are abandoned or paths that are not taken; monitoring tactical implementation against a template of expectations for signs that intervention might be appropriate; and much, much more (Schoenfeld, 1981, p. 20).

Schoenfeld found that college students are struggle with making these types of decisions when solving problems, but he had success with teaching college students to develop these behaviors in extended instruction (Schoenfeld, 1992).

Literature on solving proof-related tasks. Several studies have developed the theoretical constructs of syntactic and semantic proofs (Alcock & Inglis, 2008, 2009; Alcock & Weber, 2009; Weber & Alcock, 2004). A syntactic proof production was defined by Weber and Alcock (2004) as a proof "which is written solely by manipulating correctly stated definitions and other relevant facts in a logically permissible way" (p.210). Alcock and Weber (2009) later refined their definition using the theory of representational systems described by Goldin (1998). In this framework, a mathematical statement to be proved introduces a mathematical representative system that consists of primitive characters, e.g., words and symbols, and specific rules for manipulating those characters into permissible configurations (Goldin, 1998; Weber, 2009). Keeping the representational system framework in mind, an alternate definition of syntactic proof production is a proof that remains entirely within the representational system of the statement that requires proof (Alcock & Inglis, 2008).

In contrast, a semantic proof production is "a proof of a statement in which the prover uses instantiation(s) of the mathematical object(s) to which the statement applies to suggest and guide the formal inferences that he or she draws" (Weber & Alcock, 2004, p.210), or in terms of a representational system, a semantic proof production is a proof in which the prover does not stay entirely within the representational system of the statement that requires proof (Alcock & Inglis, 2008). Using an example of the objects mentioned in the to be proved statement to gain insight into the statement is an instance of semantic proof production. For additional insight into these two strategies view Figure 3 which originally appeared in (Alcock & Inglis, 2008).

The precise definitions of syntactic and semantic proof productions have varied slightly in the literature on the subject. Specifically, Weber (2009) defined a syntactic proof production as “when one works *predominantly* within the representation system of proof”(p. 201, *emphasis added*). However, Alcock and Inglis (2009) criticizes this choice of definition because it was too difficult to measure the amount of work completed in a particular representation in order to decide if is was predominant or not. Weber’s definition created ambiguity for deciding if a proof was syntactic or semantic (Alcock & Inglis, 2009). With either set of definitions, this language provides opportunities to compare different types of reasoning and proof writing strategies.

Studies have shown that some individuals show preference to syntactic proof production, whereas other individuals prefer semantic proof production (Alcock & Inglis, 2008; Alcock & Weber, 2009; Pinto & Tall, 2002). In one particular study, Alcock and Inglis (2008) analyzed how two doctoral students used examples while evaluating and proving conjectures. One participant used a syntactic strategy, whereas the other participant used a semantic strategy, specifically one of using examples. From this work, Alcock and Inglis concluded that expert mathematicians use different reasoning strategies when generating proofs.

As discussed earlier, individuals, whether undergraduates or professional mathematicians, may utilize several different proof strategies while constructing a proof, such as a procedural proof production, a syntactic proof production, or a semantic proof production. Each proof construction strategy provides different opportunities for learning, as shown in the theory presented by Weber (2005a). For example, a procedural proof is when an individual merely follows the steps provided by an external source, e.g., when first learning $\epsilon - \delta$ proofs, many students learn the

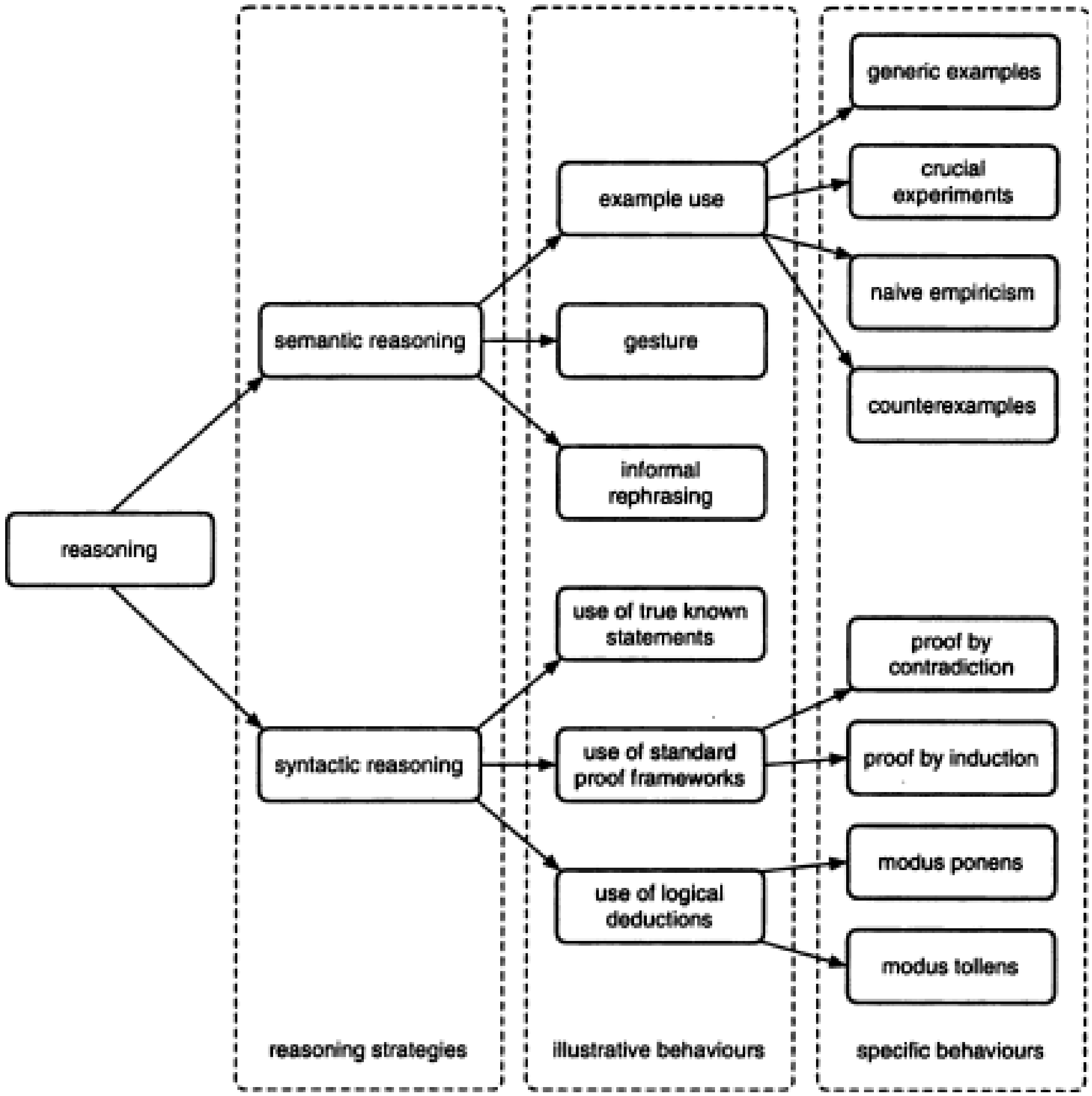

Figure 3: A model of semantic and syntactic reasoning strategies from (Alcock & Inglis, 2008).

procedures required to complete such a proof before understanding conceptually why the proof works. A procedural proof production allows a student to practice the procedures, which is a necessary component of proficiency. Similarly, syntactic proof production provides students the opportunity to use rules of inference and apply definitions and theorems, and semantic proof production provides students the opportunity to develop representations of mathematical concepts and to explore connections between concepts (Weber, 2005a). Each of the choices made by students regarding how to approach proofs provides opportunities to learn, and can change the learning environment.

Mathematicians proof reading and validation. Experience is a factor that directly affects behavior with regards to reading proofs. Weber and Mejia-Ramos (2011) began research in proof reading by interviewing several mathematicians about how and why they read proofs. The results from this study observed that when reading proofs, mathematicians claim to employ two different strategies, *zooming in* which focuses on the details of a proof line-by-line and *zooming out* which focuses on the overall structure of a proof (Weber & Mejia-Ramos, 2011). In response to this article, Inglis and Alcock (2012) used eye-tracking software to observe how long the participants spent viewing each line of a proof. They used this software to compare the habits of mathematicians and undergraduates while reading proofs, and observed that undergraduate students spent significantly longer reading formulas than the mathematicians. However, Inglis and Alcock (2012) found no evidence of a zooming out strategy in either group of participants.

The debate regarding whether or not mathematicians utilize a zooming out strategy, continued with published critiques of the eye-tracking data (Weber & Mejia-Ramos, 2013; Inglis & Alcock, 2013). After performing their own analysis, Weber and Mejia-Ramos (2013) argued that the evidence of the zooming out strategy was lost due to taking the mean over the different tasks, and because of the short length of some of the proofs. However, Inglis and Alcock (2013) countered the argument by discussing the noise found in the eye-tracking data. Overall, this series of articles argue that more details need to be determined about what is meant by a zooming-out strategy, and what are the features of a proof for which a zooming-out strategy is used (Inglis & Alcock, 2012; Weber & Mejia-Ramos, 2013; Inglis & Alcock, 2013).

Another study about expert behavior is a study documenting how mathematicians validate a proof (Weber, 2008). Validation is the process by which a mathematical argument is analyzed to determine if all of the facts are true and the inferences are applied correctly. Although validation is often considered to be a social construct, meaning acceptance from the mathematical community, Weber (2008) focused on the line-by-line analyses performed by individual mathematicians during proof validation. In the instances when a line in a proof could not be instantly accepted or rejected, the mathematicians used a variety of verification techniques: construct a formal proof, employ informal deductive reasoning, or generate examples to form an empirical argument. The use of examples varied: some strategically selected examples to observe a pattern, others used a specific example to illustrate the general properties, several attempted to find a counterexample and concluded the statement was valid when none were found, and a few used a single example and then decided the line was valid. The personal preferences, background in the subject, and ways of thinking of each mathematician impacted which strategy they chose to use during the proof validation process.

Reading proofs printed in journals can also affect personal preference. In a study of mathematicians, Weber (2010) found that professional mathematicians tend to read proofs in order to gain new insights into new ways of thinking about mathematical problems, or new ways methods of proof. The insight that mathematicians glean from reading the work of other mathematicians results in changes in the ways these mathematicians think.

Undergraduates proof reading and validation. Alcock and Weber (2005) studied the process through which 13 students validated a proof. In particular, Alcock and Weber asked if "leading prompts that focus the undergraduates' attention on relevant aspects of the proof improve the undergraduates' performance"(p. 126). Alcock and Weber found that the prompts significantly increased the number of students who made the correct claim about the validity of the proof, and thus concluded that proof validation "may be in many students' zone of proximal development" (p. 133) and that instruction in proof validation might substantially improve students' abilities in a validation task.

Selden and Selden (1995) argued that being able to unpack informal statements into formal logical statement is a necessary skill in order to correctly validate proofs. Informal statements are precise mathematical statements that use language other than those found in predicate calculus. For example, the statement "There is a function g such that $g' = f$ whenever f is continuous at each x" (Selden & Selden, 1995, p. 137) is an informal statement because the quantifiers are hidden and the implications is hidden within the word whenever. The unpacking of this statement is "$\forall f\,[(\forall x\, f \,\text{is continuous at}\, x) \rightarrow (\exists g\; g' = f)]$" (p. 137). After reviewing student work over the course of six years, it was determined that students struggle with unpacking these statements, and do not unpack them with any accuracy. As a consequence, students are unable to reliably validate arguments, because they do not understand the statements that is being proved.

Proof writing of mathematicians. The process of writing a proof is often quite different from the final proof product. In his book, Lakatos (1976) presents a series of historical case studies which illustrate the processes that mathematicians engage in while developing new mathematics. Through these case studies, Lakatos developed a theory for the process of mathematical discovery. The process begins with an initial conjecture, referred to as the *primitive conjecture*, and some idea of a proof for the conjecture. According to Lakatos, discovery occurs through the process of reconciling the proof of the conjecture with potential counterexamples.

Often the initial reaction to a counterexample is to deny that the construction is a counterexample. Lakatos (1976) dubbed this process at *monster-barring* where mathematicians respond to a counterexample by revising the included definitions in order to exclude the counterexample. Although this is sometimes appropriate, often such rules are insufficient (meaning other counterexample exist) or too broad (meaning they exclude examples). A similar process is *exception-barring* where mathematicians revise the *primitive conjecture* to exclude the counterexamples. This is different from *monster-barring* because the revision occurs with the conjecture rather than the underlying definitions. However, *exception-barring* may also lead to situations where the conjectures domain of validity is too severely restricted.

The final process described by Lakatos (1976) is the method of *proofs and refutations*. In this process a mathematician simultaneously analyzes the conjecture, the counterexample and the proof to determine any hidden assumptions in the proof. For instance, Cauchy had written a proof that the limit of a convergent sequence of continuous functions is continuous, yet Fourier had constructed counterexamples to this statement. Seidel analyzed the proof and the counterexamples, and discovered that Cauchy assumed what we now know as uniform convergence, but the counterexamples did not have this property. Ultimately, this resulted in the discovery of a new definition, and a revision of Cauchy's *primitive conjecture*. Overall, these three processes describe how mathematicians approach mathematical discovery.

Proof writing of undergraduates. The transition from computationally focused classes in the calculus sequence to more abstractly focused classes is difficult for many students. Moore (1994) studied students as they made this transition, and categorized the difficulties that students experienced. There were eight categories of difficulties: i) the prover cannot state the definitions; ii) the prover lacks intuitive understanding of the concepts; iii) the prover cannot use concept images to write a proof; iv) the prover fails to generate and use examples; v) the prover does not know how to structure a proof from a definition; vi) the prover does not understand the language and notation; vii) the prover does not know how to begin the proof; and viii) the prover does not have a sufficient concept of proof. Moore argued that a course that targets these problem areas should help students be successful through this transition, however, more integration of proof in previous courses would have an even greater impact.

In another article, (Sowder & Harel, 2003) present three case studies of undergraduate mathematics majors and their abilities to understand and produce proofs. They conclude that for a proof to be tangible to students it must be concrete, convincing and essential. Concrete means that "the proof deals with entities students conceive as mathematical objects, (i.e., objects they can handle in the same manner they handle numbers, for example)" (p. 264). Convincing means that the "students understand its underlying idea, not just each of its steps", and essential means that the "students see the need for the justification" (p. 264).

These stages are dependent on the development of the students, and cannot be determined outside of that context.

Example use of mathematicians. A recent paper by Weber, Inglis, and Mejia-Ramos (2014) explored how mathematicians truly obtain conviction about mathematical theorems and conjectures. The premise of this paper questioned the arguments that claim that mathematicians are only convinced of statement by purely deductive arguments. Weber et al. (2014) argued that this is an overstatement, because in certain instances mathematicians can obtain conviction from authoritarian and empirical sources. For instance, mathematicians will often accept published theorems from peer-reviewed journals without verifying the proof for themselves, essentially obtaining conviction from this authority. Additionally, many mathematicians believe the Goldbach conjecture to be true because of the vast empirical evidence. Weber et al. (2014) argue that we should teach students the other purposes of proof beyond obtaining conviction, since conviction can be obtained through other means.

Examples are used by mathematicians, but how are these examples constructed? Antonini (2006) sought to answer this question by conducting clinical interviews with seven mathematicians where each mathematicians was asked to construct four examples. From these interviews three distinct categories emerged: *trial and error*, *transformation*, and *analysis*. *Trial and error* is characterized by recalling known objects, and then testing whether the object has the desired characteristics. *Transformation* is characterized by taking a known object which has some of the necessary characteristics, and then performing adjustments until the resulting object has all the required characteristics. *Analysis* is characterized by considering the requires properties, and deducing what additional properties the object has to have. Eventually, this list of properties reaches a point that either a known example is evoked or an algorithm for constructing an example is determined.

Antonini (2006) observed that mathematicians often follow a process of starting with *trial and error* and then using *transformation* only when *trial and error* fails. The *analysis* technique was only used when after failing to construct an example with both the *trial and error* and *transformation* techniques. Antonini (2006) notes that the *analysis* technique is appropriate when there is a possibility that no example with the given properties exists, because the derivation of properties could lead to a proof by contradiction.

Example use of undergraduates. Alcock and Weber (2010) considered the purposes and effectiveness with which undergraduate mathematics students used examples on proving tasks. Using grounded theory, Alcock and Weber established four purposes for examples, 1) understanding a statement, 2) evaluating the truth of an assertion, 3) generating a counterexample, and 4) generating a proof. The students in this study had difficulty writing the proofs; most of the students wrote invalid proofs on the interview tasks. The students often had problems translating from the examples to the proof, in particular they had trouble formalizing their arguments.

Alcock and Simpson (2004, 2005) analyzed how a learner's beliefs about their role as learner of mathematics, and their preferences for using visual images impact their ability to make mathematical arguments in real analysis. Alcock and Simpson found that a learner's belief strongly affects their consistency between understanding examples and definitions, and the appropriate use of definitions in arguments. In particular, if a student had an internal sense of authority, meaning the student values their own thinking and judge their knowledge of definitions and concepts, then the student demonstrated consistency between examples, definitions and arguments. However, if a student relies on an external sense of authority, meaning the student relies on the validation of their instructor or other external sources to judge their level of knowledge, then the student demonstrated inconsistency between examples, definitions and arguments. With regard to students using visual or non-visual methods, Alcock and Simpson found that students who consistently visualized mathematical objects could quickly form judgments about the truthfulness of a statement. The non-visual students struggled with forming judgments and starting proofs. However, these quick judgments alone are insufficient for students to form complete arguments. The visual students need to be able to translate from the visual objects to formal definitions and then into arguments, which seemed to be a consequence of how their view their role in proving.

Buchbinder and Zaslavsky (2011) developed a series of tasks that described the construction of an example and asked the subjects to determine whether the given example was a coincidence or a consequence of the construction. The tasks were developed at the level of high school geometry and were evaluated on high-school students and teachers. From the evaluation of these tasks, Buchbinder and Zaslavsky (2011) found that the subjects performed with the most accuracy when the approached the tasks with a level of uncertainty. When the subject were uncertain if the claim was always or sometimes true, they were careful with the assumptions that they made. However, when the subjects expressed certainty in the claim, they seemed to experience overconfidence which led to the use of false subclaims and ultimately incorrect proofs. Buchbinder and Zaslavsky (2011) conclude that these types of tasks are useful for casting doubt on the ability to obtain conviction from a single example of a claim.

Pedemonte and Buchbinder (2011) studied how final year high school students used examples during a proof. The students in this study were given the definition of a triangular number as the number of dots on a filled in equilateral triangle of a certain side length, and asked to find and prove how many dots are in the triangle with n dots on a side. Some of the students focused on the patterns found in the results from calculating several triangular numbers, whereas other students focused on the patterns found in the process of calculating the numbers. Pedemonte and Buchbinder (2011) observed that the students who successfully wrote proofs focused on the process rather than the results. Additionally, Pedemonte and Buchbinder (2011) compared the argument used for forming the conjecture (the constructive argument) and the argument used for proving the conjecture (the structurant argument) through cognitive and structural lens. A student response had cognitive unity if the constructive argument and the structurant argument concerned the same concepts and some of the same logical chains. A student response had structural continuity if the constructive argument and the structurant argument used the same type of structure, such as inductive or deductive. The students who constructed proofs demonstrated cognitive unity and structurant continuity between their arguments, whereas the other students did not. This study provides insight into some of the implications from using examples, however the sample size was particularly small (only four students).

Behavior on one task can impact the conceptual knowledge gained from other topics. A particular instance of this occurred in a study by Iannone et al. (2011), where students were asked to generate examples of a particular type of function. The research team found that most students generated examples with a *trial and error* technique. Other students used a *transformation* technique where they modified known examples, or an analytic technique where the student deduced additional properties of an example. Iannone et al. (2011) theorized that the *trial and error* strategy resulted in weaker conceptual gains than the other strategies.

However, when it comes to the source of the examples used by students, Iannone et al. (2011) found that there was no significant differences between the proof productions of students who generated their own examples of a definition and those who read provided examples of the same definition. This result is contrary to other literature that supports example generation as an important pedagogical tool (Dahlberg & Housman, 1997; Moore, 1994; Watson & Mason, 2002, 2005; Weber, Porter, & Housman, 2008). In fact, Iannone et al. (2011) found that the proof productions of the example reading group was slightly higher than the proof productions of the example generating group, although the difference was not significant.

The Teaching and Learning of Mathematics

One of the primary goals of mathematics education is to develop and implement interventions that change mathematics teaching (Fukawa-Connelly, 2012a). At the undergraduate level, Speer et al. (2010) criticized that "very little empirical research has yet described and analyzed the practices of teachers of mathematics" (p. 99), even though poor undergraduate mathematics teaching is often cited as a reason students change majors away from science, technology, engineering, and mathematics fields (Seymour & Hewitt, 1997). In fact, Mejia-Ramos and Inglis (2009) conducted a literature of 102 mathematics education research papers concerning undergraduate students' experience reading, writing and understanding proofs, yet none of these papers described the instruction the students received. Although some studies have investigated instruction in proof writing since the publication of these critiques (e.g. Fukawa-Connelly, 2012a, 2012b; Mills, 2014), there is still a need for additional studies in this area.

Instructional techniques. University instructors are diversifying the instructional techniques used in their classrooms. While many classroom still utilize lecture, others use inquiry-oriented practices or inverted classroom techniques. One way to distinguish different techniques is to focus on how much of the responsibility for mastery is held by the instructor and by the students. One scale for instructional practice ranges from "pure telling" to "pure investigation" (Rasmussen & Marrongelle, 2006), where pure telling relies on the instructor for all new information, and pure investigation relies on the students exploration. Rasmussen and Marrongelle (2006) argue that well-designed inquiry-oriented courses lie close to the center of this continuum. A similar scale introduced a spectrum between noninterventionalist (the students dictate the pacing and implementation of classroom activities) to total responsibility (the instructor decides classroom activities) (McClain & Cobb, 2001). McClain and Cobb (2001) observed an instructor who tended to bounce between nonintervention and total responsibility depending on the immediate needs of the classroom, and one goal of this project was to guide the instructor towards the middle of the spectrum. For the classroom observed in this study, the instructor maintained nearly all responsibility and was closer to the pure telling side of the spectrum. The instructor delivered most of the instruction from the front of the classroom with the students sitting at tables facing the front. However, the instructor did frequently review student work, ask the students questions and responded to questions from the students which moved the instruction away from pure telling.

Lecture. Lecture has a reputation of sometimes being an ineffective means of communicating advanced mathematics ideas to students. Leron and Dubinksy (1995) argued that "the teaching of abstract algebra is a disaster and this remains true almost independently of the quality of the lectures" (p. 227). Even lectures that are masterfully executed can result in students who feel lost and confused.

One reason that students may struggle with lectures is the stereotypical description that "a typical lecture in advanced mathematics ... consists definition, theorem, proof, definition, theorem, proof, in solemn and unrelieved concatenation" (Davis & Hersh, 1981, p. 151). The formal rigor in the presentation of these proofs may hinder students from developing an intuitive understanding of why theorems are true (Hersh, 1993) and discourage them from using informal ways of understanding mathematics to produce proofs (e.g. Dreyfus, 1991).

Recent case studies have shown some lecturers provide insight beyond the definition, theorem, proof paradigm. Mills (2014) studied the types of examples provided during lectures and found that the lecturers in the case study included examples in approximately half of the proof presentations. In other case studies, the instructors provide insights into the various steps and logical structures of the proofs (Fukawa-Connelly, 2012a; Weber, 2004).

Lew et al. (2015) conducted a study where a segment of a lecture was recorded, and the students of the course were interviewed about what they learned from the lecture both from their memory and from reviewing the recording. The student participants failed to grasp the key ideas of the lecture, even after the second viewing. Although the students gained some knowledge from the lectures, they by and large missed the points the instructor was trying to make.

Suritsky and Hughes (1991) and Williams and Eggert (2002) developed a framework of how students can learn from a lecture. According to these authors, a student learning from a lecture requires a student to have four broad skills: listening (i.e., paying attention), encoding, recording the points that an instructor makes in written form (i.e., note-taking), and reviewing. However, these skills are difficult for students to enact effectively, as shown by Lew et al. (2015).

Student presentations. In some upper level mathematics courses, instructors are choosing to use students presentations of their work to progress the class (Laursen, Hassi, Kogan, & Weston, 2014). Proponents argue that putting the responsibility of instruction on the students provides opportunities for thinking critically and may improve attitudes towards mathematics (Clarke, Breed, & Fraser, 2004; Laursen et al., 2014; Smith, 2006). Fukawa-Connelly (2012b) found that the expectations for student presentations and the observations of such presentations in an abstract algebra class led to the development of norms that were useful for effective proof construction and comprehension, such as reading each other's proofs carefully and asking questions. In addition, the students began to see proof as a means of communication, rather than simply as an exercise (Fukawa-Connelly, 2012b; Knuth, 2002).

Teaching problem solving. Problem solving is difficult to teach because of its nature; a problem is a task without an immediate solution and the solution path typically requires creatively applying different techniques. One of the reasons that teaching heuristics for problem solving is difficult is that one heuristic may have implementations that are very different. For instance, Schoenfeld (1980) discusses the heuristic "exploiting simpler analogous problems" in the context of several problems where each problems used the heuristic in slightly different ways. Schoenfeld argues that "to solve a problem using this strategy, one must (a) think to use the strategy (this is nontrivial!), (b) be able to generate analogous problems which are appropriate to look at, (c) select among the analogies the appropriate one, (d) solve the analogous problem, and (e) be able to exploit either the method or the result of the analogous problem appropriately" (1980, p.796). This quote clearly illuminates how using heuristics is not simple, and teaching others to use them must address these components.

However, Schoenfeld (1980) implemented problem solving courses that explicitly teach these problem solving strategies to undergraduate students with varying mathematical backgrounds. After completing one of these courses, the students could successfully identify and implement the heuristics included in the course, even when the contexts changed, indicating significant growth. Schoenfeld argues that some applications of these heuristics may still be beyond the reach of students versed in the strategies, however this means merely that the expectations must be realistic.

Teaching proof writing. Instruction can influence the choices that students make and their preferences when solving problems, including proofs. Syntactic knowledge alone is insufficient to guarantee that a student can construct a proof; students also need strategic knowledge in order to select appropriate strategies (Weber, 2001). It is known that heuristics are difficult to teach, but that students typically do not learn them unless an attempt was made to teach them (Lester, 1994). However, some instructors do try to design the courses they teach in order to explicitly teach students strategic knowledge (Weber, 2004, 2005a).

Student often struggle with the logic required for writing accurate mathematical proofs. Epp (2003) wrote an article for instructors of upper-level undergraduate mathematics classes that specified some of the trouble with logic that these students experience. She encouraged the instructors to explicitly discuss the differences between everyday and mathematical language. Additionally, she recommended that instructors pay careful attention to the small words in proofs and definitions, such as *if*, *and*, *or* and equals (Epp, 2003, p. 896).

Mathematicians teaching transition-to-proof courses demonstrate different strategies for thinking about and writing proofs (Alcock, 2009, 2010). In the sample of two professors discussed in Alcock (2009), one professor emphasized syntactic reasoning strategies and the other professor emphasized semantic strategies. However, both professors had reasoned and justified arguments for the emphasis they chose for their students. A theoretical consequence of these different emphases in the teaching environment is that students develop different ways of thinking about proof, which then leads to different behaviors.

The theory of warranted implications, also by Weber and Alcock (2005), supports the conclusion that proof validation is within students' abilities with proper instruction. A warranted implication is a mathematical implication (i.e. if p, then q), that is both logically true and can be justified. For example the statement 'If 7 is prime, then 7 has exactly two factors, namely 1 and 7' is a warranted implication because it is logically true and we can justify the statement using the definition of a prime number. In contrast, the statement 'If 7 is prime, then 1013 is prime' is not a warranted implication because although the statement is logically true, there is no argument to justify why 7 being prime makes 1013 also be prime (Weber & Alcock, 2005, p. 34-35). Weber and Alcock argue that textbooks for transition-to-proof courses tend to emphasize only the logical evaluation of implications, not whether the implications are warranted. In addition, Weber and Alcock theorize that students will not consider whether or not implications are warranted without receiving explicit instruction in the topic, either from the instructor or the textbook, and that this skill is utilized by mathematicians when validating proofs.

In more recent work on this topic, Hodds, Alcock, and Inglis (2014) developed a module that teaches students about self-explanation techniques. Self-explanation is a reading comprehension technique where students read a segment of a text and attempt to explain the text. In the context of proof comprehension, self-explanation takes the form of explaining each line of a proof based on definitions, proof structures, logical inferences, and previous lines in the proof. Essentially, this module encourages students to find the warrants implied by the argument. Hodds et al. (2014) concluded that this module improves students' proof comprehension with effects that appear immediately and last over a period of at least three weeks.

Some instruction focuses on the processes of doing mathematics more than proof production. One instance of this was described by Larsen and Zandieh (2008), where they implemented an abstract algebra curriculum design in the theory of realistic mathematics education (RME). RME focuses on the guided reinvention of mathematics and wants the students to engage authentically in the processes of mathematics. In particular, Larsen and Zandieh (2008) showed that Lakatos' (1976) theory of mathematical discovery can be used as a heuristic for designing instruction.

Social norms are the explicit and implicit rules of classroom behavior (Rasmussen & Stephan, 2008; Fukawa-Connelly, 2012b). Norms develop over the length of a class and are negotiated by both the students and the instructor (Cobb, Stephan, & Gravemeijer, 2001). Fukawa-Connelly (2012b) studied the role of norms in an abstract algebra classroom. He concluded that norms such as students explaining and defending their work, focusing discussions on high-level ideas and asking questions of others, can aid students on proof construction and comprehension tasks.

Teaching examples. Recent work from Mills (2014) created a framework for the types of examples used in the lectures of upper-level undergraduate mathematics classrooms. The framework identified eight types of examples that were identified by the relationship to the statement of the claim and the presentation of the proof, see Figure 4. The first two types occurred before a claim was stated: start-up examples that motivate the claim, and pattern exploration which is a technique for generating a conjecture. Boundary examples occur simultaneously with the statement of the claim, and are examples that question whether the hypotheses are needed. Examples that instantiate the claim must be presented after the claim, and can occur to illuminate the claim or the proof.

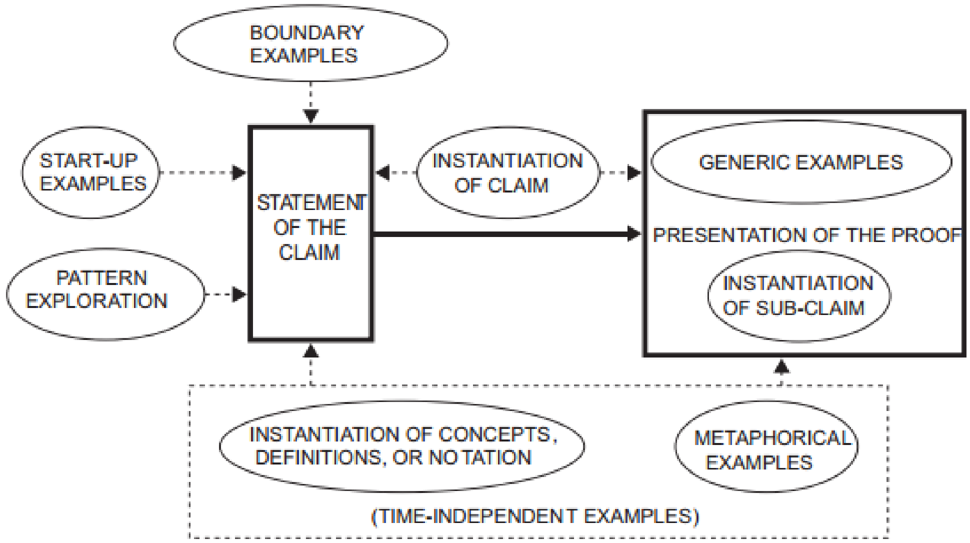

Figure 4: A framework of the example types used in the instruction of upper-level undergraduate mathematics courses.

During the proof presentation, several example types are possible. Sometimes instantiations of the claim are presented during the proof. Other times subclaims are instantiated to evaluate their truth, and to illuminate the characteristics of the subclaim. Generic examples are specific examples that have all of the characteristics of the general statement, and it is possible to prove the general statement side-by-side the verification of the example.

The final two example types occur independent of time: instantiations of concepts, definitions and notations, and metaphorical examples. Examples are a common way to improve the conceptual understanding of definitions and notations, and are frequently found throughout mathematics courses. Metaphorical examples make connections between mathematical structures, and often describe analogies.

This framework is useful for categorizing the examples used, but it is unclear which types of examples are most prevalent. Mills (2014) partially attributed the choice of the types of examples used to the personal preferences of the instructors, and to the type of course being taught.

In another study Alcock and Weber (2010) argue that some undergraduates will benefit from instruction in using examples during proofs. Alcock and Weber observed that some undergraduate students who used examples could then generate a complete proof, but other undergraduate students could not translate the informal argument they created from examples into a formal proof. One result emphasized by Alcock and Weber is that the students used examples for appropriate purposes, even when they used examples ineffectively. The argument follows that additional instruction on effective example use theoretically should result in improvements in proof writing.

CHAPTER 3

METHOD

This study is organized using qualitative research methods. Qualitative methods focus on "how people interpret their experiences. . . and what meaning they attribute to their experience" (Merriam, 2009, p. 5). Since this study is interested in the experience of an instructor teaching students to prove and the experiences of the students as they learn to prove, qualitative methods are appropriate. Specifically, this study is interested in the processes of teaching and learning the methods of proof, rather than the outcome of the instruction (Merriam, 2009).

The primary methodology for this study is that of a case study. The case for this study is a single section of a transition-to-proof course at a large university within the state of Texas. The university has a large department of mathematics, and undergraduates pursuing majors in mathematics may choose a specialization in pure mathematics, applied mathematics, or secondary teaching certification. The degree plans which require the transition-to-proof course are a major in mathematics, a major in applied mathematics, a major in mathematics for secondary teaching certification, or a second teaching field for a secondary teaching certification. Students pursuing a minor in mathematics have the option of taking the course. The selection of this course matches the definition for a case from Glesne (2011), which says that a case is "a *bounded* integrated system with working parts" (emphasis in original, p.22). A section of a course is clearly a bounded system, and the interactions between the instructor and the students are the working parts.

This study is an instrumental case study, defined by Creswell (2013) as a study where "the researcher focuses on an issue or concern, and then selects one bounded case to illustrate this issue" (p.99). This study clearly satisfies this definition, because the issue is how undergraduate provers learn to use examples to motivate their proofs, and the case consists of the students in the course.

Participants

The participants in this study are the instructor, Dr. S, and the 27 students enrolled in her course during the semester of the study. Dr. S is a respected member of the department with considerable experience teaching this transition-to-proof course, and other advanced mathematics courses. Dr. S has a reputation of being an excellent teacher, as evidenced by several awards from the department and the university for teaching excellence.

All students enrolled in the course provided consent to participate. Due to the constraints on time and resources, a nested case study design was utilized (Patton, 2002). Patton defines a nested case study as a case that is comprised of several smaller cases, which applies in this study because the section of the course is comprised of the 27 students who individually may be considered a distinct case. In an effort to understand the experiences of the students in the course, four students were selected for more detailed data collection. These students were selected during the fourth week of the semester with the following criteria: (a) the students will have demonstrated a commitment to the course by working on the assigned homework outside of class, (b) the students will have demonstrated a willingness to communicate by participating in class discussions, and (c) the students will have provided consent for participation in the study. Additionally, the students were selected using maximal variation sampling (Creswell, 2013), by varying the students' levels of academic success (indicated by a self-reported grade point average), by varying the students' mathematical preparation (indicated by self-reported grades in mathematics coursework), and by varying the students'

Table 1: The characteristics of the sampled students.

Pseudonym	Year	Major	GPA	Course Attempt
Amy	Senior	Mathematics for Secondary Teaching	2.50-2.99	3rd
Carl	Sophomore	Mathematics for Secondary Teaching	2.50-2.99	1st
Raul	Junior	Applied Mathematics and Biochemistry	3.50-4.00	1st
Mike	Senior	Mathematics and Spanish	3.00-3.49	2nd

specialization (pure, applied, secondary teaching, mathematics minor). By varying these factors, the findings have increased transferability (Merriam, 2009).

The characteristics of the four students included in the sample are presented in Table 1. These students were selected because they frequently spoke during class, both by asking the professor questions and presenting their own work on the blackboards. In addition, at least one student is pursuing each of the three specializations for a mathematics major. Unfortunately, none of the students pursuing a mathematics minor responded to the requests for interviews, so that classification of student could not be included in the sample. The students in the sample also reported a variety of different grade point averages, and different numbers of previous attempts at the course.

Data Collection

Several sources of data were used to triangulate the results (Patton, 2002; Merriam, 2009). Collecting multiple data sources is consistent with a case study methodology because a case study should present a deep understanding of the case by collecting "many forms of qualitative data, ranging from interview, to observations, to documents, to audiovisual materials" (Creswell, 2013, p. 98). The sources for this study include classroom observations, interviews with students outside of class time, and interviews with the instructor.

The primary source of data is the observations of the course. The observations are documented via field notes, audio recordings and video recordings. The field notes were taken using a smart pen, which produced an audio recording linked to the written text. The purpose of this data set is to observe the examples used by the instructor during lectures and the examples used during student presentations. Observations occurred every class period, except during exams. The researcher was a silent observer, sitting in the back of the classroom.

Interviews were conducted with the four selected students. The purpose of these interviews was to observe each student's process on proof-related tasks while working independently. Since the environment was more controlled than that of the classroom, more probing of the individual student's thought processes was possible. These interviews asked each student in the sample what he or she has learned in the course so far, and whether or not they have met the goals of the course. These interviews occurred three times during the semester: around the seventh week of the semester, the twelfth week of the semester, and the last week of the semester. The same students were interviewed at all three interview times, except the final interview could not be scheduled with Mike.

Each interview with a student had three components: a semi-structured portion addressing proof strategies and goals for the course, a task-based portion where students attempted several proof-related tasks, and a reflection on the tasks. The semi-structured portion asked the students to talk about their impressions of the course, namely what they had learned and what they thought they should be learning. The lists of questions for the interviews are found in appendices D, E, and F. The tasks for the interviews were selected from the textbook, or other studies on undergraduate proof writing (Alcock & Weber, 2010; Dahlberg & Housman, 1997; Iannone et al., 2011). The mathematical content of the questions varied over the three interviews, but matched the recent content from the course. The majority of the tasks asked the students to prove a statement, or to prove or disprove with a counterexample. However, other types of tasks such as validating proofs, generating examples and making conjectures were also included in the interviews. The students were permitted access to the textbook or their notes during the interview, and the researcher provided limited hints for clarification. After a student completed all tasks, then the students were asked to reflect on their work. Sometimes the final reflection was omitted from the interview due to poor time management.

Although the primary data source for the way the instructor uses examples is the class observations, the observations are supplemented by interviews with the instructor. The instructor was interviewed three times: prior to the semester, during week six of the semester and during week eleven. The interviews focused on the choices made during class and how those choices influenced the desired instructional goals. See appendix G for a list of questions for this semi-structured interview.

Data Analysis

As described in the previous section, the data collection includes class observations, interviews with students, and interviews with the instructor. In general, the analysis process will proceed by (a) organizing the data, (b) reviewing and memoing the data, (c) describing and classifying the data, and (d) representing the data (Creswell, 2013). Exactly how these four steps will be implemented for each type of data is described in a later section.

Case study analysis. The overall structure of this study is that of a nested case study (Patton, 2002), where each individual student included in the sample is considered a distinct case and those students are viewed together to understand the larger case of the class as a whole. In addition, the instructor observations and the interviews are distinct data sources that inform the course instruction. Then the

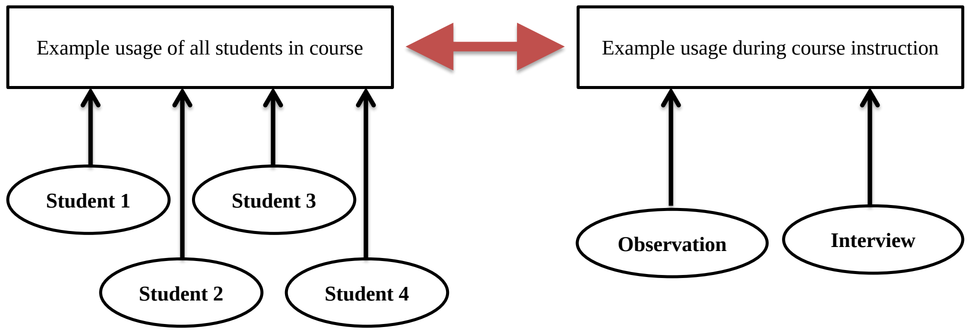

Figure 5: A model of the nested case study and case comparison design for this study.

final level of analysis will occur by comparing the two cases to one another. See figure 5 for a visual model of the structure of the design. The ovals at the bottom of the diagram represent individual cases, the boxes at the top represent the aggregate data, and the arrow represents comparing the two cases.

The analysis of the data began with a within-case analysis which means developing a holistic description of each individual (Merriam, 2009). For this study the within-case analysis means a detailed description of when and how each individual chooses to use examples. Patton (2002) describes the case study analysis process as gathering "comprehensive, systematic, and in-depth information"(p. 447) about the case. This will be accomplished by considering all of the proof constructions of each individual student, observing features of each such instance, and looking for patterns between the circumstances of the proof construction and the decision to use an example or not.

The analysis will utilize the constant comparative method developed by Glaser and Strauss (1967) for use in grounded theories. Merriam (2009) describes the constant comparative method as the process of "identifying segments in your data set that are responsive to your research questions"(p. 176), then "to compare one unit of information with the next" (p. 177), and eventually sorting all of the units into categories or themes. The process is “inductive and comparative” (Merriam, 2009, p. 175) because it occurs as the data is being collected and each new segment of data is compared and integrated into the categories as they are being developed. Once all of the data is collected the categories will be interpreted (Creswell, 2013).

Once the cases of the individual students are understood in depth, a cross-case analysis looking for commonalities and differences among the cases will occur (Merriam, 2009). By looking across the cases, it is possible “to build a general explanation that fits the individual cases”(Merriam, 2009, p. 204), and this explanation will describe the example usage during proof construction of the students in the course as a whole.

The constant comparative method of analysis will be used to interpret the instruction on example usage during the lectures. The interviews with the instructor will also be analyzed for insight on the use of examples. The themes revealed will describe the example usage in the course instruction.

The final level of analysis compares the proof construction and example usage of the students in the course with the proof construction and example usage of the course instruction. Comparing these two cases will provide insight into the impact that instruction has on the choices made by students during proof construction.

Data screening. Several types of data were collected, including video recordings, field notes of the classroom observation, and different types of interviews. The details of the data processing and analysis performed with each type of data are discussed in the paragraphs that follow.

On a daily basis, the classroom was observed and field notes were taken using a smart pen. The field notes were supplemented by video or audio recordings. The video and field notes were processed by reviewing them and identifying segments containing an example. These segments were open coded using the constant comparative method for features of the proof construction activity.

The interviews with the instructor were reviewed and transcribed in their entirety. Since the purpose of the interviews with the instructor is to explain the motivation of the instructional choices made, the interviews were coded with an open coding. These interviews supplemented and triangulated the classroom observations.

The student interviews were designed to gain a deeper understanding about which strategies the students choose to use. The interviews provided an opportunity to talk to the students one-on-one about what they think of the class, what they feel they have learned, and if they feel they are meeting the goals specified by the instructor. The portions of the interviews where the students were discussing the nature of the course and their reflections on their work were transcribed. Additionally, during the task-based portion each instance of example use was noted. The transcriptions and examples were open coded using the constant comparative method. During this coding memos were written that described different aspects of the coding, until clear themes emerged about the indicators, purpose, construction and implications.

Validity and Reliability

A well-designed qualitative research study contains several elements: an appropriate theoretical framework that guides the inquiry, especially in data collection and analysis techniques; rich descriptions of the context of the study to promote transparency; triangulation of data by using multiple sources, possibly including member checks of the findings; and sufficient time spent collecting data (Merriam, 2009). Each of these elements contribute to the validity, reliability, and transferability of the results.

A careful description of the theoretical framework that guides a study design greatly strengthens the quality of the work. Well-established frameworks often have their own criteria for evaluating the credibility of research (Patton, 2002). In addition, theoretical frameworks assist in "producing more trustworthy interpretations" (Glesne, 2011, p. 211), by supplementing findings with literature. This study will use learning theory, particularly Bandura's social cognitive theory to frame the results. This theory of learning will guide the interpretation of the results, particularly the interactions between the instructor's and the students' actions.

Rich descriptions are another important aspect of good qualitative research. Thick, rich description means "a highly descriptive, detailed presentation of the setting and in particular, the findings of a study" (Merriam, 2009, p. 227). By providing all the details of the participants, the data collection, and the data analysis, readers of the research can determine which aspects of the results transfer to other circumstances (Creswell, 2013). In addition, providing the details of the methodology also addresses the limitations of a particular study (Glesne, 2011). Specifically, high quality studies include descriptions of the participants, of the data collected, and of alternate data sources that were unavailable to researchers. By collecting detailed records, it will be possible to include rich descriptions in the presentation of this study.

Reflexivity, or addressing the position of the researcher, is another component that requires rich description in qualitative studies. Since qualitative research depends greatly on the researcher's insights, all assumptions, biases, past experiences, and theoretical orientations need to be made explicit to the readers of the research (Creswell, 2013). Reflexive description "allows the reader to better understand how the individual researcher might have arrived at the particular interpretation of the data" (Merriam, 2009, p. 219). This process of exposing a researcher's assumptions provides additional information about the context of the study and can increase transparency (Glesne, 2011). In this study, reflexivity will be incorporated into the process of analyzing the field notes. In particular, the researcher will conclude each session of data analysis by reflecting on assumptions and potential biases.

Triangulation is another technique that increases the value of qualitative research. Triangulation is the process of using a variety of data sources, analysis techniques, or theories to provide corroboration for the results (Creswell, 2013; Merriam, 2009). Patton (2002) describes several ways that triangulation can be completed. Often the simplest technique is to utilize many different types of data collection, including quantitative data, when appropriate (Patton, 2002). Alternately, when a team of investigators work on a study, it is possible to triangulate the analysis by having team members independently complete the analysis and compare their results. If a team is not possible, a variation is requesting an external audit of the data analysis processes (Patton, 2002). Triangulation is built into this study as multiple types of data are included in the design.

Member checking is a type of data collection often used to triangulate results. Member checking is the process of reviewing the initial findings with the participants to ensure accuracy and reliability (Patton, 2002). This provides an opportunity to adjust the findings to faithfully capture the perspectives of the participants (Merriam, 2009). In this study, the preliminary findings were shared with the instructor of the course to ensure an accurate portrayal of the instructor's perspectives. The students who participated in the interviews were also invited to hear the results, however none of them chose to participate.

Another aspect of good qualitative research is spending sufficient time collecting data (Creswell, 2013; Merriam, 2009). The goal in qualitative studies is to collect data until the results feel saturated, or the researcher begins "to see or hear the same things over and over again" (Merriam, 2009, p. 219). In addition, when extended time is spent in the field, trust is built between the researcher and the participants, and greater consistency between the researcher's findings and the perceptions of the participants is possible (Creswell, 2013). Since the case for this study is a one-semester course at the university, the maximum amount of time that can be spent on this study is one semester. This study is designed to maximize this time, by including observations throughout the semester and by conducting interviews at the beginning, middle, and end.

Summary

The goal of this study is to describe the circumstances in which students decide to use examples while working on proof-related tasks similar to those assigned in a transition-to-proof class, and then to compare the behavior of the students with the instructional decisions made during the course. These goals are accomplished by determining when, how and why four individual students chose to use examples (or not) during proof construction. These four cases are analyzed in depth and compared to devise a general explanation for the choices made by students in this class, considering both the personal factors and the observed behavior of the students.

Additionally, the lectures of the instructor are analyzed to determine when, how and why she chose to use examples (or not) during the course. Interviews about her instructional decisions supplement the observations. Once a general explanation for the choices in the instruction is determined, the student behavior explanation and the instruction explanation are compared to see if any relationship between the learning environment and the behavior of the students can be observed.

CHAPTER 4

RESULTS

This chapter presents the results from the study. The results beginning with additional background information about the students especially their beliefs about proof, and a description of the course design and the content. In the next section, the development of the effective example use model is described using first the data from the students and it was then revised with the data from the instructor. This description is followed by details of each aspect of the model included data from each student and the instructor.

The Student Participants and Their Beliefs About Proof

Four students out of the entire class of 27 students were selected for interviews. These students were selected because they were actively participating in class by asking and answering questions, but were not necessarily the best students. The four students were all mathematics majors, but two specialized in secondary teaching, one in pure mathematics and the final in applied mathematics. Additionally, the students had varied grade point averages, numbers of credit hours and previous attempts in the course. The following paragraphs outline the background of each of the four students, including their beliefs about proof, and their beliefs about examples.

Amy's background and general characteristics. Amy was a mathematics for secondary teaching major and a sport science minor with enough credits to be classified as a senior. She self-reported her grade point average in the range of 2.5-2.99, and this was her third attempt in the course. Amy felt that in her previous attempts in the course that she never understood proof construction; she merely memorized the proofs in her notes. Amy believed that "a proof is the background and the understanding of why it works and not just, 'it works'." However when discussing how to prove a construction is an equivalence relation, Amy said "I know I'm supposed to do these three things [to prove an equivalence relation]. I don't know why it works. I don't know how it works." So, although Amy believed that proofs should provide insight into understanding mathematical structures, she did not necessarily ascribe that characteristic to all of her proof attempts.

During the first interview, Amy classified herself as a visual and tactile learner. To help her organize her thinking, Amy would use different colored pens for different phrases in the claim to be proven. Specifically, Amy would circle each mathematical object included in the claim, and then rewrite the definition with the same colored pen. She continued to describe strategies for organization in the later interviews by saying "I write myself notes when I read through the problem just so that I don't forget to do certain things." She frequently would circle, underline or annotate parts of the problems statements, or her proof attempts during the task-based interviews.

Throughout all three interviews, Amy talked about the importance of definitions and examples. According to Amy, Dr. S would say "If you're stuck write out a definition." However, Amy said examples make definitions "more concrete and less abstract [and] a lot more helpful." Amy frequently used examples and diagrams to help improve her understanding of definitions and statements. In particular, Amy used number lines and Cartesian plane graphs to help her visualize set concepts and function properties, respectively. Amy often used the term picture instead of the term example when instantiation statements, especially on tasks involving functions.

Amy used more examples than any other student in the sample. Amy used examples 36 times over three interviews, and on 13 of the 17 proof tasks she attempted. Amy tended to focus on what she can learn from examples, and used them to expand her understanding. Unfortunately, Amy struggled using examples to generate proofs and never did this successfully, although she constructed many accurate counterexamples.

Carl's background and general characteristics. Carl was a mathematics for secondary teaching major classified as a sophomore. This was his first attempt to take the course, and he self-reported his grade point average in the range of 2.5-2.99. During the first interview, Carl defined a proof to be "the verbal illustration of a mathematical concept", and that the purpose of proof is "to define why things work, instead of just acknowledging that they do". Carl continued to describe his approach to proof as

> I try to find the right format, the right formula, that goes with the proof. I try to, I guess, bring in all the elements of the proof that I'm going to need. Am I going to need to define specific elements? Am I going to need to use certain algebraic expressions? I try to get all my pieces together before I begin a problem.
> You're going to need different variables. You're going to need different formulas or definitions or theorems. I like to try to get all those together before I start the problem so that I can analyze it as I go.

This attitude towards proof writing persisted throughout the semester. In the final interview, Carl described his proof writing process as finding a "generic outline of how it works." Related to his belief about proof writing, Carl argued that examples have limited use in writing proofs because "you can't really use specific problems to solve all of the problems that are going to be thrown at you, so it's nice to illustrate the generic things with specific examples, but you can't use specific examples to solve generic problems." Carl stated that the only purpose of an example is to have an "illustration of general principles." From an example, Carl claims he will "learn how the pieces of the general outline fit together and why they fit together." From these comments throughout the semester, it appears that Carl focuses on the step-by-step procedures of writing proofs.

Surprisingly, Carl never mentioned producing a counterexample as a purpose of constructing examples. However, when asked how he constructs examples Carl said "I just got to find an example that fills the properties of the general outline. If you fill the properties of the general outline, the example is going to fail." When asked specifically about finding a counterexample to an if-then statement, and Carl said "I'll analyze the rules for it and try to identify where it's going to fail." By the rules, Carl meant "the definition of whatever we're proving, and then I'll look for where it's going to fail. Drawing a picture or just charting numbers, and then find, 'Where's it going to fail?' And then I can go about finding why it's going to fail there." It appears that Carl's attitude towards proof extends to his understanding of examples, as indicated by repeated use of the terms 'general outline', 'properties' and 'rules'.

In the three interviews combined, Carl used 19 examples and on 12 of the 16 proof-related tasks that he attempted. Carl only attempted to construct examples on three types of tasks: tasks with the instructions *prove or disprove*, tasks that specifically asked for *example construction*, or tasks that required *making a conjecture* before solving.

Raul's background and general characteristics. Raul is a double major in applied mathematics and biochemistry and classified as a junior. This was his first attempt to take the transition-to-proof course, and he self-reported his grade point average in the range 3.5-4.0. During the first interview Raul said that “a proof is a statement that validates something.” Raul also believed that the Dr. S’s goal was for the students “to learn how to write properly” and to remember that proofs are for communicating to others.

Raul felt that examples were how he convinced himself that a statement is true. Raul claimed “you probably will not do a million of cases, but after some, you’ll end up saying ‘oh it makes sense for this one’.” Furthermore, Raul said “you should apply your previous knowledge,” and use that knowledge to create mental examples of a statement. During the first interview, Raul argued that he found examples more convincing than proofs because he lacked sufficient experience with proofs.

Raul constructed 21 examples total, and used examples on 10 of the 14 task that he attempted. Raul often used examples to clarify his understanding, and he had success connecting his examples to formal proofs.

Mike’s background and general characteristics. Mike is classified as a senior, and is a double major in pure mathematics and Spanish. This was his second attempt at the course, and he self-reported his grade point average in the 3.0-3.49 range. Mike earned a C during his first attempt at the course, but decided to take it again to develop a stronger base and to try and improve his grade. Unfortunately, Mike missed the last weeks of the semester for personal reasons, and we were unable to schedule the final interview.

Mike described proving as a process of combining a conjecture and facts that you know to form a justification for why the conjecture is true or false. From his previous attempt with the course, Mike knew that understanding the definitions and theorems are crucial for success with the material. He used his textbook extensively;

he read and reread the material because he knew that he "need[s] to know that stuff backwards and forwards." When asked how he begins proofs, Mike said that he often begins by writing down everything he knows, and sometimes he "runs tests on [the conjecture] just to see" what he can learn. Additionally, Mike talked about using examples to "see if [a conjecture] works and if it doesn't work, [to ask] why doesn't it work?" These statements seems to indicate the importance of reflection and understanding.

Mike used six examples during the two interviews he completed. These examples were used on four of the six tasks he completed. Mike worked a bit slower than the other students, so he completed fewer problems than the other students. However, Mike used examples on a high proportion of the tasks he completed.

The Instructor and the Design of the Course

The course was held twice a week for 80 minutes each session for fifteen weeks, plus a final exam in week sixteen. The instructor, Dr. S, designed the course to have three exams, the first during week seven, the second during week twelve, and a comprehensive final exam. Dr. S was interviewed before the course started, a few days before the first exam, and a few days after the second exam.

In addition to the exams, Dr. S assigned homework throughout the semester. The homework was split into two types: board work and turn-in. The students were to write their solutions to the board work problems on the chalkboards before class began. The students were not permitted to present the same problem as another student, unless the solution was different. The board work problems included tasks such as verifying an example satisfies a definition, generating an example with specified properties, and proofs at a variety of difficulty levels. The turn-in problems were almost exclusively proofs, and were usually due the class period following the board work presentations on the same section. Although Dr. S (and consequently I) use the phrase students presentations, Dr. S was the one who reviewed the problems verbally. The students did not talk about their work, except if Dr. S asked the student or the class for clarification or the motivation for the response.

This teacher focused approach to the student presentations was by design. Dr. S includes these board work presentations as a way to increase the amount of feedback that she can give to the students, and to provide the students time in class to get assistance on difficult problems. Specifically, Dr. S frequently reminded the students that it was acceptable to present solutions that they suspected were incorrect, because the presentation were an opportunity to get individualized feedback on their work. A few students would present solutions that included an indicator of confusion, e.g. a star and a note saying "I'm stuck here". Nearly every time this occurred, Dr. S would reach that point in the proof and ask the student why they chose this problem in order to reiterate that the board work presentation is the time to make those mistakes and that the students got credit for presenting whether or not the problems were correct.

A typical class day began with Dr. S discussing the board work presentations. She would go to each written solution and review the solution with the class. Dr. S would often ask questions of the student who wrote the solution, and would edit the solution until it was correct. After this, Dr. S would often lead a discussion of the turn-in homework that was returned. Although they did not always discuss the turn-in homework, Dr. S would be sure to discuss it if many students solutions had the same errors. In the remaining time, Dr. S would lecture on new material.

The content of the first two weeks focused on statements and logic: propositions, conditionals, biconditionals, and quantifiers. The content included truth tables, but the emphasis was on translating the text and understanding the logic. The course then moved into basic proof techniques, namely direct proof, proof by contraposition, and proof by contradiction. For these two weeks, the class proved a variety of statements about the integers, such as x is odd if and only if x^2 is odd. Dr. S emphasized the set-up for these proofs, by discussing how the words *and*, *or*, and *not* in the statement can suggest proof techniques. The remaining weeks before the first exam were focused on sets, set operations, families of sets, and mathematical induction. Most of the proofs about sets required inclusion or double-inclusion.

During the two classes after the first exam, the content focused on strong induction, the well-ordering principle and basic combinatorics (e.g., multiplication principle, permutations, combinations). The course then proceeded with relations, equivalence relations, partitions, and ordering relations. The final two class periods before the second exam focused on functions.

The final few weeks of the semester focused on additional properties of functions including the definitions of one-to-one and onto. Included in this was discussions about sequences and basic topology of the real numbers. Additionally, the students saw an introduction to the cardinalities of infinite sets including Cantor diagonal theorem. However, the terms countable and uncountable were carefully avoided.

Throughout the semester, Dr. S focused on the students understanding the details of the definitions and statements. For instance, she required that the students explicitly state that the integers are closed over addition, when they used that property. Due to this emphasis, Dr. S used examples to instantiate and explore new definitions, to understand a statement, and to reveal logical inconsistencies.

The students presented problems on the board throughout the course, especially during the first weeks of the semester and the weeks immediately after each test. After the first test, Dr. S threatened to give the students a pop-quiz if they did not present solutions to at least ten tasks every day. Although there were usually fewer than ten presentations the threatened pop-quiz never materialized. However, Dr. S intended the student presentations as opportunities for the students to see additional examples and for Dr. S to provide feedback on example and proof construction for the students.

A Model of Effective Example Use

Developing a model of effective example use began with the resource management literature, which focuses on knowing which strategy to try, whether the strategy was implemented properly, and what conclusions can be drawn after implementation (Carlson & Bloom, 2005; Polya, 1957; Schoenfeld, 1992). The focus of resource management is on "not just what you know; it's how, when, and whether you use it" (Schoenfeld, 1992, p. 60). Applying this theory to using examples, four phases were developed: indicators of examples, purposes of examples, construction of examples, and implications of examples (see Figure 6). Indicators are what inspires the use of examples, in other words, when is the strategy appropriate. Purposes are the original stated goal of the examples, in other words, what can be learned from using examples. Construction refers to knowing the techniques of construction and the accuracy of specific constructions, in other words, how examples are constructed and used. Implications are the next steps after the example is constructed, in other words, what is learned or concluded from using examples. The purposes and implications are very similar, the key difference is the purpose is what the prover intended to achieve from the strategy and the implication is what the prover actually did achieve from the strategy.

These phases were developed by looking at the theories of research management and by looking at instances of effective and ineffective example use. A discussion of some of the instances are describe in the next subsections. Then, the

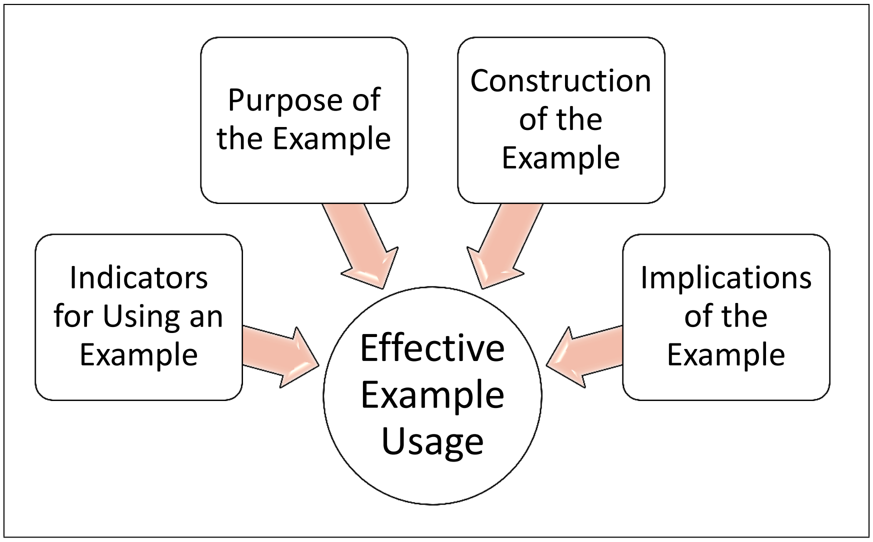

Figure 6: A model of effective example usage that includes four phases.

categories within each phase were developed and refined through open coding of the examples used by the students, the examples used by the instructor and the statement made by the instructor regarding example use. These categories are discussed in later sections.

Effective example use from the students. The student used examples effectively on several tasks during the interviews. The primary observation of these instances is that the student successfully transverses all four phases of the framework when an example is used effectively.

During the third interview, Carl successfully used examples on the task below.

> **Task** (Interview 3, Question 4a) A real valued function is called fine if it has a zero at each integer. Prove or disprove: The product of a fine function and another function is fine.

After reading this task, Carl decided to try an example because the task said *prove or disprove*. On previous tasks with this language, Carl had stated "[I will] start with a counterexample because you only need one," and his behavior on this task

Assume $f(x)$ is a fine function, and $g(x)$ is any other function. Then $f(x) \cdot g(x) = h(x)$ for some function $h(x)$. Then for every $x \in \mathbb{Z}$ $f(x) = 0$, so $h(x) = 0$. Thus $h(x)$ is fine.

Figure 7: Carl concluded his work on question 4b of interview 3 with this proof.

appears consistent with those statements. Then Carl constructed the example $x^2 \cdot \sin x$, acknowledging that he meant the transformation of sine which has it's zeros on the integers. From this example, Carl observed that a product function has a zero anywhere that one of the functions also has a zero. He then continued to write the proof seen in Figure 7.

Carl's process on this task went through all four phases. First he recognized that examples are useful for deciding the truthfulness of tasks with the language *prove or disprove*. Then he acknowledged the purpose of example construction by saying that he knew that an example would provide some evidence of truthfulness, and that a counterexample would disprove the statement. At this point, Carl used the *transformation* technique to construct an example. In the last phase, Carl decided the statement was true, and deduced properties from his example that was the insight required for writing an accurate proof.

Task (Interview 1, Question 1) Let a, b, and c be natural numbers and $\gcd(a, b) = d$. Prove that a divides b if and only if $d = a$.

Raul used examples effectively during the first interview on the greatest common divisor task shown above. Raul spent considerable time contemplating how to set-up a proof to this problem. He began by separating the biconditional statement into "$\triangle$ if $a|b$, then $a = d$" and "$\square$ if $a = d$, then $a|b$," where the $\triangle$ and $\square$ were merely symbols to identify the two statements.

Raul then decided to attempt a proof of the $\triangle$ implication, but spent more than a minute deciding what he should assume evidenced by his statement, “I'm seeing what one to assume. To assume the antecedent and go forward, or doing a contrapositive.” Raul decided to try a proof by contraposition, and wrote “Assume $a \neq d$. Then base on $\underline{\text{gcd}}$ $a \nmid b$ saying that d is $\gcd(a, b)$.” Raul explained his written statement by saying that “since d is the greatest common divisor of a and b, and a is not equal to d, *um*, then a will not divide b. But then, that's probably also wrong, because if we let $a = 2$ and $b = 4$, then the greatest common divisor will be 2, which is the same [as a, indicated by drawing an arrow].” When questioned later in the interview, Raul said that he picked the variable values randomly for the examples he constructed in the interview. At this point, Raul realized that he did not know how to proceed, and I suggested that he attempt the other direction.

When beginning the proof of the □ conditional, Raul wrote “assume $a = d$. Then it means that a will be the gcd of b too.” Raul commented that this statement “was the same as here” as he circled the example he constructed earlier. From this observation and by looking up the definition of the greatest common divisor, Raul concluded that a must divide b.

Raul's experience also has several phases. First, he recognized that he did not fully understand the structures that he was attempting to use in his explanation. Then he utilized a *trial and error* technique to construct an accurate example, as evidenced by his claim of choosing the numbers randomly. Although Raul was unable to complete a proof of the △ conditional, the example did reveal the hole in his argument. Additionally, Raul's consideration of the example contributed to his observation that $a = d$ means a has all of the properties of the $\gcd(a, b)$, and ultimately to a proof of the □ conditional.

The students were also successful finding counterexamples to statements. For instance, Amy and Carl both successfully found counterexamples on question 3a of interview 2, seen below.

> **Task** (Interview 2, Question 3a) Let $\mathcal{A} = \{A_\alpha : \alpha \in \Delta\}$ be a family of sets, $\Delta \neq \emptyset$, and B be a set. For each statement either prove the statement is true or give a counterexample.
>
> $$B - \left(\bigcap_{\alpha\in\Delta} A_\alpha\right) = \bigcap_{\alpha\in\Delta}(B - A_\alpha)$$

Amy and Carl both began by stating that they would try to find an example or counterexample in order to evaluate whether the statement was true or false. Both students wrote a small family of finite sets for $\mathcal{A}$ and then established a set B that would produce a counterexample. After they established that the construction was a counterexample, both students concluded the statement was false.

Within this study, when students correctly constructed counterexamples, or were able to make additional progress on their proofs, they went through all four phases: indicator, purpose, accurate construction, and implications. The exact nature of these phases varied, for instance, the phrase *prove or disprove* indicated the use of examples to Amy and Carl, but Raul thought to use examples after he experienced confusion. Overall, the students used examples effectively on 20 tasks, out of 53 attempted tasks. The students attempted to use examples on 40 of the tasks, so which means half of the examples they used were not effective. In all 20 of these instances, at least one phase of the process was incorrect or incomplete.

Ineffective example use from the students. Effectively using examples requires all four phases of the process to be completed. When one or more phases of the process failed, then the student used examples ineffectively.

One instance of ineffective example use occurred when Carl attempted question 4 in the second interview, which is quoted below. When attempting this problem, Carl wrote $\frac{4!}{2!}$ and $\frac{4!}{(4-3)!3!}$ on his paper and nothing else. Carl expressed confusion and that the combination formula was the only formula he had in his notes. Although Carl expressed confusion and the task actually provided an example to begin with, Carl chose not to work with the outcomes from this example. He did not seem to recognize the indicators that working with the example could be a useful strategy for gaining understanding and completing the task.

> **Task** (Interview 2, Question 4) The number of four-digit numbers that can be formed using exactly the digits $1, 3, 3, 7$ is less than $4!$, because the two 3's are indistinguishable. Prove that the number of permutations of n objects, m of which are alike, is $\frac{n!}{m!}$. Generalize to the case when m_1 are alike and m_2 others are alike.

Amy used examples effectively and ineffectively on the same problem, during the first interview.

> **Task** (Interview 1, Question 1) Let a, b, and c be natural numbers and $\gcd(a, b) = d$. Prove that a divides b if and only if $d = a$.

Amy set up the bijection proof by separating the implication and its converse into two separate proofs, and translated some of the assumptions into variables. Amy then performed some algebraic manipulations that resulted in the equation $dl = am$, from which she hoped to conclude $d = a$. Amy question whether this subclaim was true, and then generated the counterexample $1 \cdot 15 = 3 \cdot 5$. Amy realized that the subclaim was not true and that she needed to try another method for the proof.

While trying to remedy her proof on this question, Amy commented, "I feel like I have to use this somewhere," where *this* is a condition of the greatest common divisor definition "if $c|a$ and $c|b$, then $c|d$." Yet, she never used the statement. At the end of the interview, I called Amy's attention to the fact that she made the statement, but did not use the fact, and her response was:

> Amy: I just didn't know where ... Because this is a proposition and not a statement. These are direct statements, this is what it is and this is what it is. I didn't know how to translate this into something I could work with. If c divides a well then cool but what if it doesn't? Then does this whole thing fall apart or can I even use it?
>
> SH: I don't know.
>
> Amy: But I felt like this was a big 'if' and I was scared to work with it because I felt like it was a proof inside of a proof and I didn't know where to go with that.
>
> SH: You weren't confident that you had anything that satisfied that 'if.'
>
> Amy: Yeah, even though I mean I don't know I guess it's like given that all of this is true, if this is true, but I didn't know where ... I guess I just didn't know where the c lived. Like if I could just pull it out of thin air and say, "Well, because the GCD is d then there's some integer that divides d and divides a and divides b."

In this single problem, Amy used examples successfully to evaluate the truth of a statement, but did not use examples to gain understanding. Amy could have used an example to help her understand the final condition of the greatest common divisor definition, or to locate the problem in her proof. Although it is possible that Amy would not have gained anything from considering an example, it is also possible that she would have.

Carl, Mike and Raul all used examples ineffectively on question 2 of the first interview. This task is a true statement, although it includes the phrase *prove or disprove* in the directions.

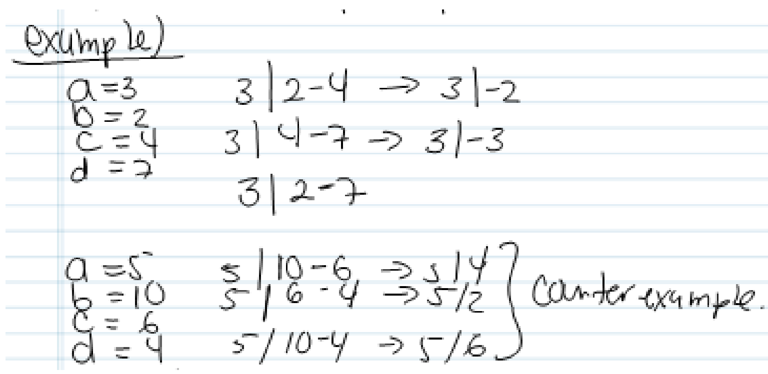

Figure 8: The non-example constructions generated by Raul in question 2 of interview 1.

Task (Interview 1, Question 2) Prove or disprove: For integers a, b, c and d, if $a|(b-c)$ and $a|(c-d)$, then $a|(b-d)$.

Raul had difficulty in the construction phase of the problem. He was utilizing a *trial and error* technique, but did not completely understand the characteristics required for a construction to be an example or a counterexample. As seen in Figure 8, Raul randomly picked four numbers and then tested whether they satisfied the hypotheses and conclusion. Raul abandoned the first construction because of the negative values. For the second construction, Raul labeled this as a counterexample because it did not satisfy the hypotheses or the conclusion. Raul concluded the statement was false because he had constructed a counterexample. At this point in the semester, Raul clearly did not have sufficient knowledge of example and counterexample constructions to be able to accurately complete the construction phase of effective example usages.

Mike was challenged with counterexample construction, although his inaccuracy was more subtle. Mike began the problem by stating that he wanted to seek a counterexample in order to disprove the statement. Mike thought about what

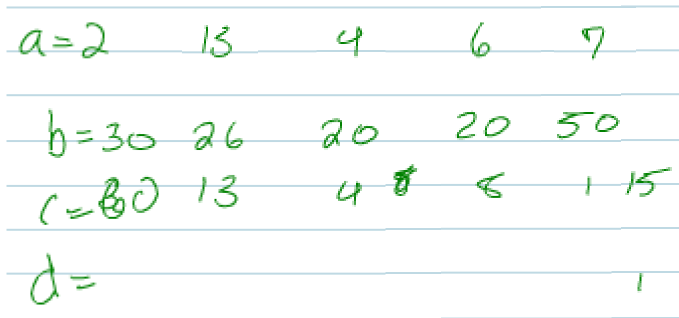

Figure 9: The list of examples generated by Carl in question 2 of interview 1. He was unable to draw a correct implication from this work.

value would be best to use for a, the divisor in the problem. He started with $a = 2$, but quickly decided to use $a = 1$ because "one divides everything." Mike did not realize that his choice of using $a = 1$ makes it impossible to achieve his desired goal of finding a counterexample. Mike did conclude that the statement was true and then wrote an accurate proof, but he never realized the contradiction between his stated goal and his example construction. In this particular case, Mike's example construction did not impede his success with the problem, however, in the next problem which was a false statement he was unable to construct a counterexample.

Carl experienced difficulty determining the implications from his examples on this task. Carl recognized that the task said *prove or disprove* and that searching for a counterexample could be a productive use of his time. Carl began by selected a, and then b and c, so that the first hypothesis was met. Then he would try to pick a d value that satisfied the second hypothesis, but not the conclusion. You can observe in Figure 9 that he could never find a value for d that would work. Additionally, Carl used a variety of numbers for a, including primes and composite values. However, Carl was very reluctant to state that the statement was true, and was unsuccessful in writing the proof. At the end of the interview, Carl asked me about the proof of this statement and when I told him it was a consequence of just adding the statements together he responded that he observed the pattern in his series of examples, but that "he didn't think he could do that." Carl was completely successful in the first three phases of effective example use on this problem, but experienced a break down in the implication phase.

The above evidence shows that a failure within any phase will results in not utilizing examples effectively. It is important to note that examples are not always needed when writing proofs of statements, and often when examples are used they occur in the scratch work or the margins and not in the final proof product. For instance, there were a few questions in these interviews where examples were not used by the students, and the students wrote successful proofs. However, there were also instances where the students did not successfully solve the problem, and working with examples might have been a fruitful endeavor.

Summary. Effectively using examples on proof-related tasks involve a four phase process of thinking to use examples, deciding why you might want to use examples, constructing the examples and finally analyzing what can be learned from these examples. If any of these phases are missing or incorrect, the example attempt will often not be useful to the prover. Additionally, the phases must be aligned, meaning each phase needs to coordinate with the previous attempts. If a prover decides to attempt an example or counterexample construction because the task says *prove or disprove*, then their purpose should be to decide truthfulness. If they choose another purpose, they may not be successful in achieving their goal.

In the next sections, each of the four phases will be discussed in great detail. In each section, themes of the different types of thoughts or behaviors that occur in each segment will be discussed and consolidated into categories. The categories were established through the analysis of the student interviews, and then revised through the analysis of the observations of the class and the interviews with Dr. S. As such, each of these sections are organized by presenting data for each student individually, and then comparing the students to one other. This is followed by how the phase was modeled in the classroom, and discussed but not modeled by the instructor. Finally, connections between the students and the instruction are discussed within each phase.

Indicators for Examples

The categories for the indicators phase were established by analyzing the components that inspire the prover to use examples as a strategy for solving the proof-related task. All of the instances of example use from the students were analyzed for the indicators, which were developed into categories. Then the lectures and the interviews with the instructor were analyzed to expand and refine the categories. The categories that emerged fit into two classifications: the language of the task, and seeking insight.

Several types of tasks indicate the appropriateness of examples. The most frequent type of task (in the interviews and the homework) are those that include the directions *prove or disprove*. Since most claims are universally quantified, the negations of these statements are existentially quantified and thus statements can be disproved by constructing a single counterexample. Thus, the phrase *prove or disprove* usually indicates that the prover should attempt to find a counterexample, or a few examples that might suggest the statement is true.

Other tasks ask the prover to construct an example with specific properties or to verify that a particular construction satisfy some properties. Only one task of this nature appeared in the interviews, but they were fairly common on the homework. In particular, the students had a series of homework problems that asked them to construct a relation that was transitive, but not symmetric or reflexive, or a relation that was symmetric and reflexive, but not transitive, and other similar problems.

Finally, a few conjecturing or generalizing tasks were asked in the interviews and on the homework. These tasks do not have a consistent phrasing, but often the words *generalize*, *conjecture* or asks a question with the words "what conclusion" or "what conditions". In these tasks, examples can be used to determine a pattern or the necessary properties.

In the subsequent sections, relevant evidence for these categories are discussed for each student individually, and then all of the students are compared. Not every category and subcategory is discussed for each student, but the most relevant evidence is included.

Amy. Amy was prompted to use examples by the language of the task and by monitoring her thinking. In ten instances Amy used an example to gain insight into a statement or definition, after expressing confusion, doubt, or an inability to move forward. In the remaining 26 instances of example use, the decision to use examples was prompted by the language of the task; 17 examples on *prove or disprove* tasks, four examples on *make a conjecture* tasks, two examples on *example generation or verification* tasks, and three examples on a task that asked the students to *generalize*.

One of Amy's strengths was her ability to monitor her own thinking and to recognize when she did not fully understand a definition or statement. While working on the problems in the third interview, Amy demonstrated that she tends to think about functions in terms of graphs on a Cartesian plane. Amy began question 1, which asks whether the composition of two decreasing functions defined on an interval is increasing or decreasing, by sketching a Cartesian plane, setting a specific interval and sketching decreasing functions. However, Amy began to doubt her intuition and reached for the textbook to verify the definitions of increasing and

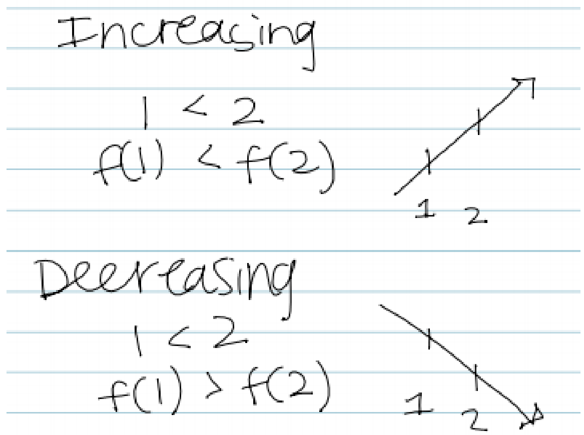

Figure 10: The examples drawn by Amy to understand the definitions of increasing and decreasing during question 3 of interview 3. She knew to draw these examples from monitoring her thinking and recognizing that she did not fully remember the definitions of increasing and decreasing functions.

decreasing. Then she used quick sketches of a function on an interval to understand the definitions of increasing and decreasing (see Figure 10).

After Amy felt comfortable with the definitions, she began working with specific functions. She used simple polynomial functions as examples, and found appropriate counterexamples to both statements. In this instance, Amy recognized that she did not fully know the definitions of increasing and decreasing and she used example graphs, example equations, intuition, and her textbook to improve her understanding.

Although Amy usually used examples when she worked on a task with the phrase *prove or disprove* she was the only student who did not use an example on question 2 of the first interview, quoted below.

> **Task** (Interview 1, Question 2) Prove or disprove: For integers a, b, c and d, if $a|(b-c)$ and $a|(c-d)$, then $a|(b-d)$.

Amy read the task and immediately wrote a correct proof to the statement. When asked about why she did not choose to construct an example even though it contained the language *prove or disprove*, Amy replied that since "[the question] involved like subtraction and like addition and subtraction is like low level. I can manipulate it a little easier. ... I feel, like in #2 to get what I want I can manipulate what I'm given, and I just kind of saw that, like I have c's in both, but I have $-c$ and $+c$, so I thought I could manipulate that easier". Later, Amy confessed that she did not consider whether or not the statement was true, she simply saw how she could manipulate the symbols to reach the conclusion.

This was the only instance that Amy attempted a task in which the language indicates that examples would be useful and Amy did not construct an example or counterexample. In addition, this particular question stands out because Amy was the only student whose proof process was completely accurate on this question.

Carl. Carl was prompted into constructing examples by the phrase *prove or disprove* in the directions of the task for 15 of the 19 examples he constructed. The remaining four examples were instigated by tasks that asked the prover to *verify or construct an example* (twice), to *make a conjecture* (once) or to *generalize* an example (once). Carl never seemed inspired to construct an example in order to gain insight or because he was confused.

During the first interview on one of the *prove or disprove* tasks, Carl stated that "I want to start with a counterexample because you only need one." On nearly all of the *prove or disprove* tasks given, Carl began by constructing examples or counterexamples. The two exceptions to this occurred when Carl did not seem to understand the concepts enough to construct an example.

Task (Interview 3, Question 2) Suppose f is an increasing function. Prove or disprove that there is no real number c that is a global maximum for f.

One instance of this confusion occurred on questions 1 and 2 of the third interview. Carl appeared to have serious trouble using the notations of sets and functions correctly. Carl wrote $x \subset I$, $g(x) \subset I$, $(a, b) \in I$, $x \in a, b$, $x \in f$ and $x + 1 \in f$, where I in an interval and f is a real-valued function. It seems that Carl was confusing $\subset$ and $\in$, but he also did not seem to understand how functions relate to sets.

Additionally, Carl is peculiar in that he never used an example to clarify his thinking. As discussed earlier in the ineffective example use segment, Carl struggled with question 4 of the second interview, repeated below. This task is atypical,

> **Task** (Interview 2, Question 4) The number of four-digit numbers that can be formed using exactly the digits $1, 3, 3, 7$ is less than $4!$, because the two 3's are indistinguishable. Prove that the number of permutations of n objects, m of which are alike, is $\frac{n!}{m!}$. Generalize to the case when m_1 are alike and m_2 others are alike.

because it includes an example within the task language. Carl's attempt to work with this example was limited to observing that $n = 4$ and $m = 2$ and writing $\frac{4!}{2!}$ and $\frac{4!}{(4-3)!3!}$ on his paper. Carl argued that "he could not get any traction" on this problems and quickly gave up. He did not even attempt to write any of the outcomes of the example provided, and demonstrated little understanding of the permutation and combination formulas. He did not seem to recognize that working with this example could improve his understanding.

Raul. Raul was generally prompted to use examples from the language of the task and by monitoring his thinking. Specifically, Raul used examples eight times to gain insight when he was confused by a statement. Raul also used seven examples for *prove or disprove* statements, three examples to *make conjectures* and three to *construct or verify an example.*

During the final interview, Raul had trouble with the definitions for increasing and decreasing functions. After reading the task below, Raul realized that he did not recall the formal definition for increasing and decreasing. He considered the functions $f(x) = -2x$ and $g(x) = -2x^2$ to try and deduce the formal definitions from his informal understanding. These examples were insufficient for Raul to recall the definitions, but they did help him make sense of the definitions after he looked them up in the textbook.

Raul also used examples to generate conjectures. On the task below, Raul constructed a few examples to try and determine the necessary conditions for g. He tried several different examples, including $g(x) = x + 2$ and he eventually deduced the correct condition of $g(0) = 0$. Although, Raul did not write a formal proof, he did articulate that "if x is an element of the integers, then $f(x)$ should equal to 0" and that to conclude that $g(f(x)) = 0$ for $x \in \mathbb{Z}$, then $g(0)$ must equal 0.

> **Task** (Interview 3, Question 5) Let $f : \mathbb{R} \to \mathbb{R}$ be a fine function. If $g : \mathbb{R} \to \mathbb{R}$ is a function, such that $g \circ f$ is a fine function, what conditions must g satisfy? Does g have to be a fine function as well? What is the weakest condition that g must satisfy to ensure the composition is fine?

Mike. Mike used examples both to gain insight and due to task language. In the six examples Mike constructed, three were used in response to confusion with regards to the definitions or statements, and three were used on *prove or disprove* tasks.

Mike decided to use examples to gain insight on the very first question he attempted. Mike realized that he did not fully understand the formal definition of the greatest common divisor, so he attempted to construct examples. The first example he constructed had the values $a = 5, b = 12$ and $d = 5$, but Mike realized that these values do not satisfy the condition that $\gcd(a, b) = d$. Mike then considered $a = 6, b = 12$ and $d = 6$. Although this example does satisfy the conditions, Mike was not able to connect the example to the proof and he did not write an accurate proof for either direction of the biconditional statement.

Comparison of the students. The students generally decided to use examples after reading the language of the task (see the language subtotal line of Table 2). The majority of the tasks (14 out of 19) had language that encouraged example use, so this result is not surprising.

Table 2: This table summarizes the indicators that inspired the students to use examples in the interviews. The language subtotal is the sum of the categories in the bottom section.

Indicators	Amy	Carl	Raul	Mike	Total
Seeking Insight	10	0	8	3	21
Language Subtotal	26	19	13	3	61
Prove or Disprove	17	15	7	3	42
Make a Conjecture	4	1	3	0	8
Construct/Verify Example	2	2	3	0	7
Generalize	3	1	0	0	4

On the *prove or disprove* tasks, the students usually began with constructing either an example or a counterexample before attempting a proof. Two exceptions to this occurred during the first interviews with Amy and Mike. On question 2 Amy immediately wrote a perfect proof without considering any example, and on question 3 she constructed a counterexample. In contrast, Mike started with examples on question 2, but started with a proof on question 3. The tasks are shown below.

> **Task** (Interview 1, Question 2) Provide either a proof or a counterexample for the following statement. For integers a, b, c, and d, if a divides $b - c$ and a divides $c - d$, then a divides $b - d$.
>
> Note: This claim is true.

Task (Interview 1, Question 3) Provide either a proof or a counterexample for the following statement. For integers a, b, and c, if a divides bc, then either a divides b or a divides c.

Note: This claim is false.

What makes this situation interesting is the reasons that Amy and Mike gave for their decisions to use or not use examples. Amy thought question 2 was easier than question 3 (due to the operations chosen, but perhaps she remembered it from the homework), and so she decided to try a proof first. Amy was less confident about question 3, so she tried examples. Conversely, Mike thought question 2 was more complicated and question 3 was more simple. He did not specify why he thought it was more complicated. Both students provided the same reason for starting with a examples instead of a proof, but different interpretations of the "hard one."

Both Amy and Mike knew that constructing examples and counterexamples could help them evaluate the truth of the claims. However, they both relied on their intuition to determine whether or not to pursue the strategy. Intuition is a valuable tool in mathematical research for determining which questions to pursue and potential answers, but it is a tool that may lead to errors.

Another interesting comparison is the number of examples used to gain insight or clarification on definitions and statements. Approximately 25-50% of the examples used by Amy, Raul and Mike were indicated by confusion and a desire for clarity. This starkly contrasts with Carl who never used examples for this purpose. It is difficult to say what caused this difference, however, Carl seemed to focus on the rules and seemed to always want the efficient answer. As a consequence, I infer that he did not see the point in constructing an example unless there was a clear advantage to do so, such as finding a counterexample that disproves the statement.

In fact, other than the very first interview, Carl never constructed more than one example for a particular task. The other students do not appear to share this attitude.

Dr. S's modeling. With the students we saw two categories of indicators: self-monitoring to determine moments of confusion, and finding key words phrases in the task language. In terms of task language, four phrases seemed to prompt the students to use examples: *prove or disprove*, *make a conjecture*, *construct or verify an example*, and *generalize*. Additional categories emerged when the instruction was analyzed for the presence of indicators. In particular, the instructor indicated that new definitions are an inidcator to use examples and that examples are useful when questioning assumptions.

Dr. S wanted the students to construct examples when they encounter a new definition, when they do not know what to do, on prove or disprove tasks, or when they form a conjecture. These categories are slightly different from what was observed from the students. In particular, encountering a new definition was not observed with the students, but was the most frequent during the lecture.

Dr. S provided the class over 100 examples or non-examples of definitions during the the course. Several of these examples had the purpose of introducing notation. Dr. S used more examples to instantiate and clarify definitions than all of the other types of examples combined.

When introducing a new definition, Dr. S nearly always provided examples without commenting on the behavior. However, when Dr. S defined an ordering relation, she stopped to ask the students "what comes next?" to which the students replied "an example." In this instance, she explicitly instructed the students that they should immediately attempt to construct examples of new definitions in order to improve their understanding. This was the only time this behavior was explicitly explained; although Dr. S modeled it the entire semester. After this discussion, Dr. S provided "two big examples" of ordering relations and continued with the class.

Dr. S frequently responded to student questions by providing an example. In one instance, the class was proving the task below.

Task Prove that x is odd if and only if x^2 is odd.

When beginning the proof of the converse, Carl asked "why can't we assume that $x^2 = 2(2j^2 + 2j) + 1$?", which was the final line from the initial direction. In response, Dr. S asked Carl to consider $m = 3$. Dr. S continued saying $3 = 2(1) + 1$, but $1 \neq 2j^2 + 2j$, and that we cannot know that every odd number has that form. Dr. S used an example to show why Carl's proposed set-up for the proof would not work.

In addition to answering student questions with examples, Dr. S frequently demonstrated how to use examples to question assumptions. For instance, when proving "if $xy = 0$, then $x = 0$ or $y = 0$" she asked the students if the statement were true in domains other than the real numbers. She provided counterexamples in clock arithmetic and matrix multiplication to show the importance of the given hypotheses. Later in the semester, when defining δ neighborhoods of a point in $\mathbb{R}$ and $\mathbb{R}^2$, Dr. S talked briefly about alternative metrics, such as the taxicab metrics. She did this to help the reveal the students' underlying assumptions about various structures and facts, and to provide quick introduction of some of the ideas that the students will see in their advanced courses such as abstract algebra and topology.

Additionally, Dr. S encouraged the students to question the statements themselves. In one lecture towards the end of the semester, Dr. S presented a theorem about the images and preimages of sets under a function. The theorem consisted of four statements, three of which were statements of set-equality and one which was merely set-inclusion. Immediately after writing the statements, Dr. S asked the class “why is in not equality?” and to realize there is probably a counterexample to equality. Dr. S ultimately provided the class two counterexamples to the equality statement. The following class period, a student presented a similar statement with set-inclusion. Dr. S instructed the class that when a text only provides a set-inclusion statement to them, they should question whether or not the opposite inclusion holds. Although in this particular case the opposite inclusion was the next question on the homework set, it was an opportunity to reiterate the importance of questioning assumptions and results.

Dr. S did model how different types of tasks could indicate examples use. In two circumstances during an early semester lecture, she asked the students to evaluate if a quantified statement was true or false, and if false to produce a counterexample, see the tasks below. Both of these questions are false, and Dr. S (with input from the students) produced counterexamples. However, since the statements were false, there was no need to produce additional examples or a proof. These questions introduced counterexamples and the phrase *prove or disprove* to the students, although the true introduction occurred in the homework set.

Task

$$\forall x \in \mathbb{R}, \frac{x^2 - 9}{x - 3} = x + 3$$

Task Every continuous function has a (closed from) antiderivative.

The next class period Dr. S continued to introduce the idea of using examples to evaluate the truth of claims that include the phrase *prove or disprove.* Dr. S asked the students if $A \backslash B = A \cap B^C$ is true or false (where B^C is the complement of set B). To approach this question, Dr. S constructed the sets $U = \{1, 2, 3, 4, 5, 6, 7\}$, $A = \{1, 2, 3, 4\}$, and $B = \{2, 4\}$, as an instantiation of the statement. After this example, Dr. S led the class through a double inclusion proof of the statement. However, she never referred to the specific example during the double inclusion proof. Dr. S demonstrated determining the probable truth of an assertion, but did not show the students how "examples can provide ideas."

Dr. S rarely modeled using examples to discover patterns and conjecture. In two instances near the beginning of the semester, students modeled this behavior on board work questions. In both questions the task asked the student to find the union and the intersection of a family of sets indexed on the natural numbers. On the first problem, the student listed the elements of the first three sets in order to determine the union and intersection of the family. On the second problem, the student drew the sets on a number line to the same end.

Another instance of conjecturing occurred later in the semester, over several lectures. When defining equivalence relation, equivalence classes and partitions, Dr. S used the same examples, including modular arithmetic with several moduli, the rational numbers, and complex numbers with the same distance from the origin. After proving the equivalence classes of five different equivalence relations form partitions, Dr. S directed the students to consider the pattern. Dr. S proceeded to prove that given any equivalence relation on set A, the set of equivalence classes form a partition of A.

Overall, Dr. S instructed the students to consider examples when they see new definitions, experience confusion, have a question, when a task asks them to *prove or disprove* and when a task asks them to conjecture. One classroom activity incorporated several of these indications. After defining various set operations such as union, intersection and complement, Dr. S specified the following sets:

- U =students in this classroom
- A =students wearing glasses

- B =students wearing gray
- C =students wearing a jacket/sweater
- D =students wearing tennis shoes

Dr. S then proceeded to ask the students a series of questions to practice with set operations including:

- Stand up if you belong to set A.
- Stand up if you belong to set D^C.
- Stand up if you belong to set $A \cup B$
- Stand up if you belong to set $(A \cup B) \cap (A \cup C)$
- Stand up if you belong to set $A \cup (B \cap C)$

The series of questions ended with the last two questions listed. With these two questions, Dr. S had the students look around and realize that these two sets appear to have the same elements. This led to the class forming the conjecture $(A \cup B) \cap (A \cup C) = A \cup (B \cap C)$, and to Dr. S proving this conjecture for the students. Through this activity, the Dr. S demonstrated many of the indicators of example use, including: exploring new definitions thoroughly, looking for patterns, and forming conjectures.

In summary, Dr. S modeled the behavior of instantiating new definitions nearly every day of the semester. She used examples to explore the nuances of the definitions. Additionally, Dr. S modeled using examples when a task includes the language *prove or disprove*, or *make a conjecture*. Finally, she modeled using examples to answer questions and to provide clarity.

Dr. S's discussion. During the interview with Dr. S during week seven of the semester, she said she wanted the students to consider examples "anytime they're stuck and don't know what else to do. Anytime they don't thoroughly understand the definition. Anytime they see a new definition whether they think they thoroughly understand it or not." The few times the students complained that they had an insufficient number of examples for definitions and other problems, Dr. S reminded them to do more board work questions. Dr. S selected the homework questions so that the board work contained the questions that involved expanding example spaces.

Dr. S also wanted the students to use examples when they were confused or did not know how to proceed. During week four, Dr. S told the class

> Somebody told me in office hours today, I get the set-up, but I'm stuck in the middle.' I am very happy with that, because what I've taught you is the set-up. The middle is get out your scrap paper, and try and bunch of stuff until something works. There's not a clean algorithm for a lot of the middle. And that's ok. I don't mind that. But you just need to not stop. Keep trying, keep working, keep going.

Later during the second interview, Dr. S repeated this sentiment when she relayed the advice she provides her students "if you don't know where to go now, get out your scrap paper and try two or three examples and now see if you see the general pattern." At this point, Dr. S wanted the students to construct examples as an attempt to understand the statements of their proofs.

This attitude was reiterated in the final interview when Dr. S said she wants her students to consider examples "when they're confused. When they're confused, when they don't understand something, when they see something new I want them to immediately try to poke a hole in it by taking examples, seeing how it works in the example." Throughout the semester, Dr. S modeled this by responding to student questions with a clarifying example. Additionally, Dr. S claimed that "I am trying to every so often in class say okay, this isn't working, why is this not working? Let's try an example." She did not model this behavior very often in class, especially not in her lectures. This is mostly because she was always prepared for her lectures and consequently did not have problems that were not working. However, when she corrected the students' presentations, Dr. S frequently utilized examples, either by constructing examples or counterexamples to explain the student's work, or by correcting the examples and counterexamples constructed by the students.

Dr. S instructed the students to generate examples and counterexamples on tasks with the instructions *prove or disprove*. She repeatedly stressed that the students knew that only one counterexample is needed to disprove most statements. She also encouraged the students to "try a couple examples, [because] sometimes examples can provide ideas," however she did not model this behavior in the lecture.

In summary, Dr. S discussed using examples to gain insight when a prover is confused or does not know how to proceed. She also talked about some of the language indicators without always modeling them, especially the phrase *prove or disprove*.

Comparing the instruction and the students. The indicator phase for using examples has three categories: understanding new definitions and concepts, clarification and insight when uncertain, and the language of a proof-related task. Several different types of task language may inspire the use of examples, such as the directions *prove or disprove*, *make a conjecture*, *generalize* and *construct* an example with specific conditions. Additionally, a statement that contains an existence quantifier may also indicate that finding an example may be appropriate.

The majority of the examples included in the lecture were instantiations of new definitions. With every new definition, Dr. S would instantiate and explore the nuances of the concept with examples. Although she modeled this behavior for the students, she did not talk about the importance of doing this to understand the ideas. In the interview just prior to the first test, Dr. S said that she wants the students to consider examples, "anytime they're stuck and don't know what else to do. Anytime they don't thoroughly understand the definition. Anytime they see a new definition whether they think they thoroughly understand it or not." Dr. S intended this behavior to extend not only to definitions, but also to theorem statements. She did instantiate several theorem statements during the course. Sometime the instantiations came before the theorem to motivate it and other times after to clarify. However, instantiating definitions was far more common than instantiating theorem statements, largely due to the simplicity of many of the theorems proven in the course.

All four sampled students mentioned the importance of examples for understanding definitions and statements when talking about strategies for success in the class, so the students seemed familiar with this indicator. The only new definitions that the students encountered during the interviews were the definitions of fine function and periodic function during the third interview. The three students who conducted this interview all constructed at least one example of a fine function, as this was one of the interview tasks. None of the students constructed a function and proved it satisfied the definition of periodic, although Amy at least thought a little about $y = \sin x$, and found the that the period is 2π. Although the students recognized that they should construct examples for new definitions, they did not always do so.

During class, Dr. S talked about using examples for "lost cause" maneuvers, i.e. a technique to try after exhausting other ideas. Similarly, she told the students that examples can reveal patterns to help them "figure out" the middle of proofs. During the first interview, several students used examples to gain insight on the greatest common divisor question. The students collectively seemed to lack confidence with the greatest common divisor definition. Raul, the only student who did not look up the definition of greatest common divisor, began with outlining the two directions of the bidirectional proof. After this, Raul was not sure how to proceed, so he constructed a simple example of the statement. This provided him with sufficient insight to prove one direction of the proof, as discussed extensively in the earlier results section about Raul.

However, not all of the students chose to use examples when they did not know what else to do. In particular, Carl did not construct a single example after experiencing confusion. Even in circumstance where he did not know what to do, he often chose to give up or move onto another problem rather than spending time on examples. One instance of this occurred when working on a generalization of a combinatorial problem (Interview 2, Question 4). The question even included a specific example in the statement, but Carl did not take the time to explore the given example. Carl was the only student who never used examples as a consequence of self-monitoring.

The final indicator of using examples is the language of the task. Dr. S told the students "to disprove, find one counterexample", a fact that the students frequently recited during the interviews. Following up this advice, Dr. S instructed the students "if [the directions] say prove or disprove, try a couple examples. Sometime they can provide ideas." The students constructed examples and all of the *prove or disprove* tasks they attempted, with only two exceptions. As such, it is clear that the students recognized that the *prove or disprove* language is an indicator to use examples.

A few other language indicators included the terms *generalize* or *conclude*. Both of these are phrases included on questions that ask for a conjecture, and most of the students attempted to use examples to try and see patterns and reach conclusions. This behavior was modeled very limitedly in the course, but it did occur. The students seemed to understand how examples could help with conjecturing tasks.

Overall, the students seemed to know when using examples would be useful. This corresponds to the instruction given about the technique. The students may have benefited from additional experiences with conjecturing tasks and existence proofs, but they seemed to recognize the indicators that were emphasized in class.

Summary of the indicators. Ultimately three categories were developed to indicate example use: 1) seeking insight, 2) exploring new definitions, and 3) task language. Task language is subdivided into four subcategories: i) prove or disprove, ii) make a conjecture, iii) construct or verify an example, and iv) generalize.

Purposes of Examples

The categories for the purposes phase were established by analyzing the stated intentions for using examples on proof-related tasks. The students did not always verbally state their intended purpose, and in these cases the purpose was inferred from the indicator, the implication and other statements. In general, the students used examples to gain a better understanding of a statement or definition, to evaluate whether the task statement is true or false, to make a conjecture and occasionally to generate the argument for a proof.

Amy. Amy primarily used examples to understand a statement and to evaluate the truth of a statement. Specifically, she used 15 examples for the purpose of understanding a statement, 16 examples to evaluate whether a statement is true or false, three examples to make conjectures, and two for generating a proof.

Task (Interview 1, Question 1) Let a, b, and c be natural numbers and gcd $(a, b) = d$. Prove that a divides b if and only if $d = a$.

Amy used examples to evaluate whether a claim was true or false on several problems. One instance of this occurred in the first interview on question 1 (shown above). Amy set up the bijection proof by separating the implication and its converse into two separate proofs, and translated some of the assumptions into variables. Amy then performed some algebraic manipulations that resulted in the equation $dl = am$, from which she hoped to conclude $d = a$. Amy question whether this subclaim was true, and then generated the counterexample $1 \cdot 15 = 3 \cdot 5$. Amy realized that the subclaim was not true and that she needed to try another method for the proof. Although Amy did not complete a proof of this statement, she did successfully use the example to evaluate the truth of her claim.

Amy used examples to gain understanding frequently, especially if she did not understand the definitions fully. When Amy realized that she did not fully understand the definitions of increasing and decreasing, she looked up the definitions and drew the sketches shown in Figure 11 to gain understanding.

> **Task** (Interview 3, Question 4a) A real valued function is called fine if it has a zero at each integer. Prove or disprove: The product of a fine function and another function is fine.

Amy behaved similarly on question 4a (above). Amy had minimal understanding of the fine function definition; in particular she had recognized that if she transformed $\sin x$ or $\tan x$ that those would both be fine functions but had considered no other examples. Additionally, Amy did not fully understand the definition of periodic. After reading the definition of periodic, Amy decided to gain some additional incite by applying the definition of a periodic function to $f(x) = \sin x$. Amy did successfully argue that $f(x) = \sin x$ is a periodic function,

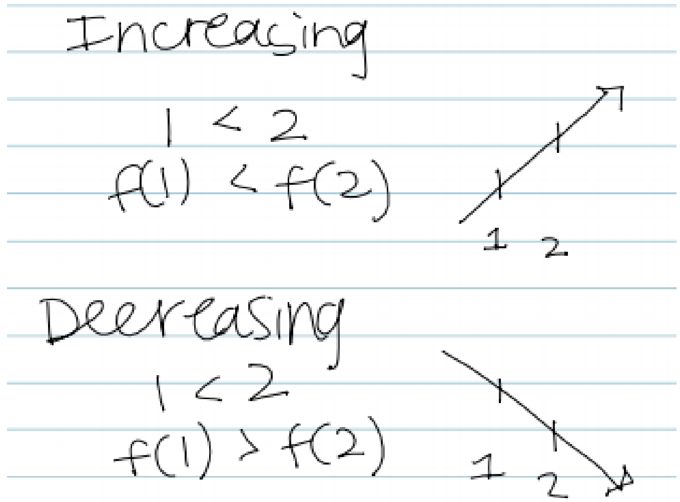

Figure 11: Amy used these examples for the purpose of understanding the definitions of increasing and decreasing.

and appeared to gain additional insight into the definition. Unfortunately, Amy forgot what she was trying to prove, which ultimately led to her having problems at the implication phase.

Carl. Carl primarily used examples to evaluate whether a statement is true or false; out of the 19 examples Carl used, 15 of them were constructed with the purpose of evaluating the truth of the statement. Carl constructed examples to understand a statement three times, and to make a conjecture once.

> **Task** (Interview 3, Question 4b) Prove or give a counterexample: The product of a fine function and any other function is a fine function.

One instance of evaluating the truth occurred during the final interview on question 4b. Carl immediately considered $(\sin x) \cdot x^2$, although he knew that $\sin x$ was not truly a fine function. From this example, Carl concluded that the statement is true. He then realized that multiplying by zero always makes the product zero, and then quickly generalized to an accurate proof (see Figure 12). Although Carl

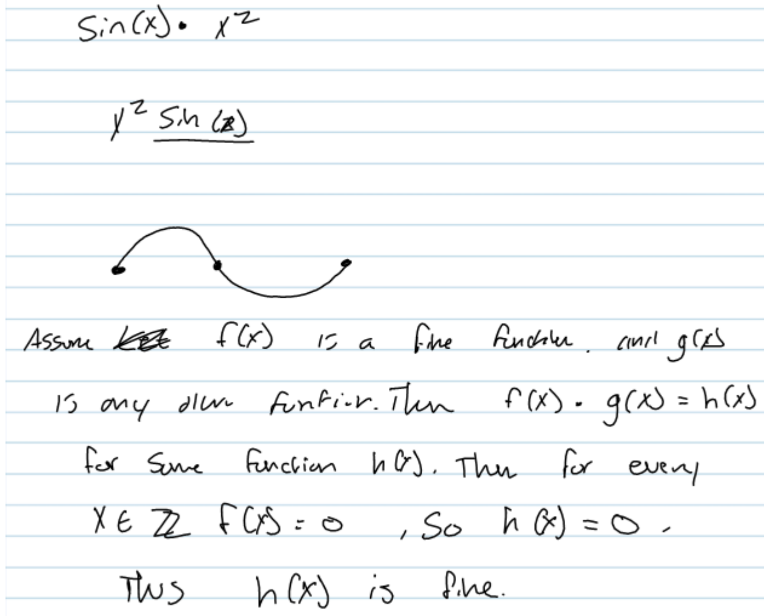

Figure 12: Carl's proof that the product of a fine function and another function is fine.

ultimately used the properties from the example to generate a proof, his intended purpose was to evaluate whether or not the statement is true.

Raul. Raul used examples for multiple purposes including 10 examples for understanding a statement, nine for evaluating whether a statement is true or false, one for making a conjecture and one for generating a proof.

During the first two interviews, Raul used examples primarily to evaluate if a statement is true or false. In fact, he only used one example during these interviews for the purpose of understanding the statement, and one for generating a proof. However, in the third interview, Raul primarily used examples to understand statements. This is partially due to the tasks selected in the third interview; Raul used examples to clarify his understanding of the definitions of increasing function, decreasing function, and fine function.

Task (Interview 2, Question 4) The number of four-digit numbers that can be formed using exactly the digits $1, 3, 3, 7$ is less than $4!$, because the two 3's are indistinguishable. Prove that the number of permutations of n objects, m of which are alike, is $\frac{n!}{m!}$. Generalize to the case when m_1 are alike and m_2 others are alike.

Raul used examples to understand the statement and to generate a proof on question 4 of the second interview. Raul began by exploring the example provided in the task statement for the purpose of understanding the statement. However, Raul felt that this example was insufficient for him to be able to see the pattern. He decided to construct another example, $1, 3, 3, 3$, in order to learn more about the pattern and to try and write a proof.

Mike. Mike used examples for the purposes of understanding a statement and to evaluate whether a statement is true or false. In particular, Mike constructed three examples for each of these purposes. For the three examples that he constructed for the purpose of understanding a statement, Mike decided to use an example because he realized that he did not fully understand the definitions and that he needed additional insight. Additionally, the three examples that Mike constructed for the purpose of evaluating the truth were inspired by tasks that include the directions *prove or disprove*.

Comparison of the students. The evidence suggests that the students were aware of the various purposes that examples are used for during proof writing and related tasks. In particular, the sampled students could all recite the role of counterexamples in disproving statements, and they attempted to construct examples or counterexamples on nearly every *prove or disprove* task in the interviews. Deciding the truthfulness of statements was the most frequent purpose for which the students used examples, however, this may be a consequence of the proportion of *prove or disprove* tasks included in the interviews (see Appendices D, E and F to see the language of the tasks).

Table 3: This table provides a summary of the purposes for which the students used examples during the interviews

Purpose	Amy	Carl	Raul	Mike	Total
Understand a Statement	15	3	10	3	31
Evaluate True or False	16	15	9	3	43
Make a Conjecture	3	1	1	0	5
Generate a Proof	2	0	1	0	3

Other than Carl, the students constructed examples for the purpose of understanding statements with great frequency (see Table 3). Amy, Raul and Mike all seemed to appreciate how specific examples can provide insight into generalized concepts. Carl's reluctance to use examples for this purpose is unusual compared to the other students. This may be a consequence of Carl's step-by-step approach to proof writing. Although, Carl described the purpose of proof writing as knowing why things are true, his process of proof writing tended to focus on following rules instead of increasing understanding.

> **Task** (Interview 2, Question 4) The number of four-digit numbers that can be formed using exactly the digits $1, 3, 3, 7$ is less than $4!$, because the two 3's are indistinguishable. Prove that the number of permutations of n objects, m of which are alike, is $\frac{n!}{m!}$. Generalize to the case when m_1 are alike and m_2 others are alike.

Although some of the students mentioned making a conjecture and generating a proof as purposes for examples, these purposes were rare. The only examples that were constructed for generating a proof occurred on the combinatorics question. Amy and Raul both began with the question included in the task, but then expanded to similar examples in order to gain more information in the hopes of finding the pattern and generating a proof. Although neither one was able to do so, they both stated generating a proof as the purpose for these additional examples.

Dr. S's modeling. Dr. S modeled using examples for the purposes of understanding statements and definitions, making connections between different definitions, revealing logical inconsistencies in arguments, and forming conjectures. Throughout the semester, many new definitions were introduced. Dr. S provided examples that instantiated the definitions given in the lecture. She also frequently provided non-examples and boundary examples that revealed the sufficiency and necessity of the hypotheses. For instance, after defining a prime number as an integer a such that $a|(b \cdot c)$ implies $a|b$ or $a|c$, Dr. S wrote $4|(6 \cdot 10)$, but $4 \nmid 6$ and $4 \nmid 10$. This shows that 4 is not a prime number. Dr. S followed this with the example $4|(6 \cdot 12)$, $4 \nmid 6$ and $4|12$, which shows that a composite number may satisfy the hypotheses for *certain values*, but only a prime number will satisfy the condition *for all* such values.

Dr. S often used examples to clarify the nuances of definitions, statements and notations. In one instance, Dr. S observed the students struggling with the notation for relations. Previously, Dr. S introduced multiple notations for relations including $(x, y) \in \mathcal{R}$, $x\mathcal{R}y$, $\mathcal{R}(x) = y$, listing the entire set of ordered pairs that belong to $\mathcal{R}$, and drawing a directional graph. She informed the students that the context determines the usefulness of the notations, where the context means the number of ordered pairs in the relation and the elements in the domain and codomain sets. Later in the semester, Dr. S informed the class that "there's one example that I want to do, primarily for the notation involved, which tells you that the notation is going to be [pause] a little bit messy." Dr. S continued to say there was a homework problem that utilized the notation from this example. She then provided the example of relation T on $\mathbb{R} \times \mathbb{R}$, where $(a, b)T(c, d)$ if and only if

$d - b = c - a$. This relation is notationally complicated because it is a relation on a set of ordered pairs, so the elements of the relation are ordered pairs of ordered pairs. However, Dr. S connected the symbolic notation to the graphical interpretation in the Cartesian plane, specifically that two points of $\mathbb{R} \times \mathbb{R}$ are related if the slope between them equals one. From this, Dr. S led the class through a discussion showing that T is an equivalence relation where each equivalence class is a line of slope 1. The representation in the Cartesian plane may have helped the students realize that an ordered pair can be viewed as a single object rather than two distinct objects that happen to be paired. Dr. S clearly intended this example to provide additional understanding of the notation for relations.

The definitions became significantly more complicated throughout the semester. Most of the examples early in the semester could be constructed without proof; we generally do not expect a prover to show that 5 is prime when using it as an example of a prime number. Later in the semester the definitions were generally sufficiently complex that a construction required proof to show that it is actually an example. For instance, proving a family of sets is a partition requires three short proofs, one for each condition of the definition. This increase in complexity contributed to a change in the homework tasks assigned. When the definitions became more complex, more homework tasks asked for proofs that a given construction is an example. Whereas with the easier content, more questions focused on proving statements that used the definitions in an abstract sense. This observation is important because it shows that the role of example instantiation did change through the semester, and was content dependent.

Another purpose of examples modeled by Dr. S is the instantiation of connections between definitions. This was observed the day before the second exam, when Dr. S introduced a family of functions $A_\alpha = \{(x, y) \in \mathbb{R} \times \mathbb{R} | y = \alpha - x^2\}$, and posed the task, "show that $\mathcal{A} = \{A_\alpha | \alpha \in \mathbb{R}$ is a partition of $\mathbb{R} \times \mathbb{R}\}$." Dr. S then guided the class through viewing a function from $\mathbb{R}$ to $\mathbb{R}$ as a set on $\mathbb{R} \times \mathbb{R}$, and then to view the family of functions A_α as a partition of the set $\mathbb{R} \times \mathbb{R}$. During this episode, Dr. S not only reviewed several of the difficult definitions that the students could expect on their exam, but she showed how the elements of this family of functions can be viewed as functions, as relations, and as sets, and how the entire family is a partition and ultimately as the equivalence classes for an equivalence relation.

Another instance when Dr. S used examples to make connections occurred when she introduced the definition of a one-to-one function. Dr. S connected the formal definition of a one-to-one function from generic sets A to B to the horizontal line test given in precalculus to determine if real-valued functions are one-to-one. She did this by drawing a graph of a real-valued one-to-one function and showing the students how the definitions applied. Additionally, Dr. S discussed how the horizontal line test is more than sufficient when the function is on the real numbers, but does not work well on functions on "wacky sets" that cannot be drawn. Through this example, Dr. S situated the new material within the existing knowledge of the students and exposed the connections.

Although the primary purpose for which Dr. S used examples was to instantiation definitions, she also repeatedly used examples to reveal logical inconsistencies in proofs. The first instance of this occurred when she warned the students about assuming the conclusion in their proofs, when she wrote an argument that $-1 = 1$, see Figure 13. Although this is a sample argument, it is an argument of about a specific fact. Additionally, Dr. S produced this argument twice during class, and discussed it during one of the interviews, so she viewed this as an important example. This argument reveals the fallacy in this line of reasoning

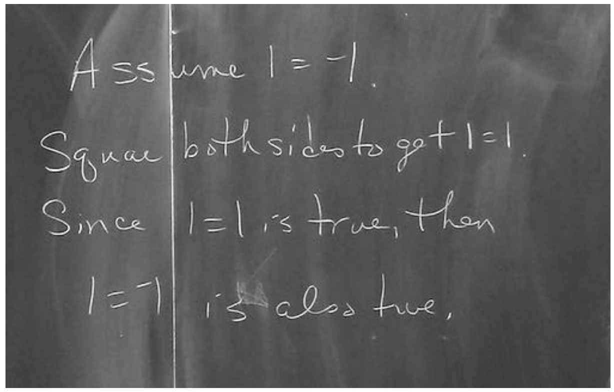

Figure 13: Dr. S's warning about assuming the conclusion.

because the conclusion is clearly false. When the students presented this type of reasoning later in the semester, Dr. S would remind them about this argument.

Another instance when Dr. S used examples to reveal logical inconsistency occurred when several students wrote the following on a homework assignment:

$$x + y = \frac{a}{b}, \quad a, b \in \mathbb{Z}, b \neq 0$$

$$\text{If} \quad a = c + d, \quad \text{then} \quad x + y = \frac{c+d}{b} = \frac{c}{b} + \frac{d}{b}$$

$$\text{Thus,} \quad x = \frac{c}{b} \quad \text{and} \quad y = \frac{d}{b}.$$

When reviewing this problem, Dr. S provided the example $1 + 4 = 2 + 3$, to show the students that their conclusion is not logically valid.

Dr. S introduced counterexamples to statements, particularly when the students included false statements in the proof presentations. During one presentation, a student failed to address the assumption of non-negativity in the proof that $\sqrt{xy} \leq \frac{x+y}{2}$. As such, Dr. S asked the students if they could produce a counterexample when the non-negative condition is removed. The class produced a few numbers, and Dr. S wrote $x = -3$ and $y = -3$ on the board. Dr. S followed this with a discussion of the "power in counterexamples". In particular, she addressed

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Figure 14: Dr. S provided this boundary example. The product of matrices is a boundary example of the zero-product theorem.

the future secondary students in the classroom about the common algebra mistake $\sqrt{x^2 + y^2} = x + y$. Dr. S asked these students "how do you get [your students] to remember this for five years?", and her answer was give them a counterexample. Dr. S ended the exchange by writing the counterexample $x = 3$, and $y = 4$, because $\sqrt{3^2 + 4^2} = 5$, but $3 + 7$. This episode reveals that Dr. S wanted the students to appreciate the role of counterexample, and hoped that this level of reasoning will be something the students take with them as they continue their studies and careers.

On several occasions, Dr. S used examples to express the importance of the hypotheses. Early in the semester, she proved the following theorem as an example of proving by cases.

Theorem Suppose $x, y \in \mathbb{R}$. If $xy = 0$, then $x = 0$ or $y = 0$.

Immediately after writing the theorem statement on the board, Dr. S mentioned that the theorem is not true in all worlds. Specifically, she talked about clock arithmetic and specifically that $3 \cdot 4$ is equivalent to 0 in clock arithmetic. A few weeks later, when reviewing homework with the class, Dr. S was concerned about a students use of this theorem without acknowledging that the objects involved were real numbers. At this point, Dr. S provided the class with another world with a counterexample, namely the product of matrices, see Figure 14. Dr. S told the students to be cognizant of the hypotheses of theorems, especially for those statements that they instinctively know.

Finally, Dr. S demonstrated using a collection of examples to form a conjecture, particularly by generalizing patterns. An episode illustrating this purpose began with Dr. S asking the students if they recalled the binomial theorem. When they did not, Dr. S led the students through the expansions of $(a+b)^1$, $(a+b)^2$, $(a+b)^3$ and $(a+b)^4$, where in each iteration used the previous iteration as a foundation, *i.e.*, $(a+b)^4 = (a+b)(a+b)^3$. Due to time constraints, Dr. S did not take the students through the entire process of generating the conjecture, but provided the students with the closed form of the statement. Dr. S did not prove the entire statement, but noted that the iterative form permits a proof by induction, and instructed the students that the full details of the proof was available in the textbook. Although Dr. S did not fully model the process of using examples to form conjectures, she did model how examples can provide the opportunity to seek patterns and which can lead to increased understanding of the ideas.

In summary, throughout the semester Dr. S modeled using examples primarily to help the students understand the subtleties of definitions and statements. This included examples, non-examples, boundary examples, and making connections. She also used examples to reveal problems or potential problems with the students' reasoning, which included talking about the power of examples and how counterexamples can provide a mechanism for remembering information.

Dr. S's discussion. Although she did not model it, Dr. S did discuss that examples can be used to help generate proofs. At the beginning of the semester, Dr. S taught the students how to set-up direct proofs, proofs by contraposition, and proofs by contradiction. The focus of this instruction was on the first and final lines of the proofs, not the middle. When the students complained that they could not complete the middle of proofs, Dr. S replied that "the middle is get out your scrap paper, and try and bunch of stuff until something works." Dr. S continued this discussion by saying that examples can be used to find patterns or properties and can be useful for figuring out the middle.

Throughout the semester, Dr. S discussed how examples could provide insight when the students were stuck in their work. Dr. S did not model this use of examples, but in several instances she mentioned that examples can reveal the patterns or properties that are needed for the proof. To further her case, Dr. S mentioned that plugging an example (or counterexample) into a incomplete proof can reveal problem areas in a proof attempt.

Comparing the instruction and the students. During the course the instructor modeled using examples to understand statements, to evaluate the truth of the assertions, and to reveal logical inconsistencies. She also discussed using examples to gain insight into problems and to find patterns or properties of the objects in the statements. The students used examples for similar purposes, specifically evaluating the truth of assertions, understanding statements, and generating proofs. The students did not use examples to reveal logical inconsistencies, however, this is likely a consequence of the tasks given to the students.

One way that Dr. S encouraged the students to use examples to reveal logical inconsistencies was by verifying mathematical statements using concrete numbers. Specifically, when translating the statement a divides b for a and b integers, Dr. S suggested the students take the equation $ak = b$ for some integer k and substitute a simple example such as $a = 2$ and $b = 6$, to get $2k = 6$. She recommended this to the students because in her previous experience teaching this course, she has seen several students write $bk = a$ instead of $ak = b$. However, by plugging in numbers, it becomes clear that $6k = 2$ in inconsistent with k being an integer, but $2k = 6$ is not. Dr. S explicitly mentioned using examples in this manner twice in the lecture and presumably more often on homework and exams.

With the repetition of this statement, it seems possible that the students would utilize it during their work, especially when the definition of divides was used repeatedly. Yet, none of the sample students exhibited this behavior during the interviews. All four students applied the definition of divides correctly without considering an example (at least without considering an example in an observable way). Their apparent comfort with the definitions likely contributed to the students not utilizing examples in this manner.

Although Dr. S frequently modeled using examples to reveal logical inconsistencies, the students did not talk about this purpose of examples or utilize it in their work. Since this use of examples typically resulted in a counterexample, the students may not have realized the difference between revealing inconsistencies and seeking counterexamples to decide truthfulness. Also, the students were constructing their own arguments, whereas, Dr. S typically used examples to reveal inconsistencies when correcting or discussing the students' work. Due to these reasons, it is not surprising that the students did not fully comprehend this purpose of using examples on proof-related tasks.

Dr. S modeled constructing examples for the purpose of understanding a definition or statement many times throughout the semester. Nearly every new definition was instantiated with one or more example. Additionally, Dr. S would occasionally use examples to clarify the theorems to be proved. Sometimes the example instantiated the statement, but often it was a boundary example which showed why the hypotheses were needed.

The students imitated this behavior, especially later in the semester. The students recognized the value of such behavior, such as when Raul said "I'll try to see what I understand first, and try to make that into numbers." Amy also argued that "seeing the definition as ... more concrete and less abstract is a lot more helpful." Amy and Raul both seemed to appreciate the insight they gleaned from examples, and frequently used examples for the purpose of understanding the statements more fully. Mike and Carl did not explicitly mention this purpose of examples, and they used examples for this purpose far less frequently.

Another important purpose of examples is to evaluate the truth of a statement and to construct a counterexample if the statement is false. Dr. S modeled this behavior on several occasions, especially at the beginning of the semester. In additional instances, Dr. S talked about the role of counterexamples and how they are used to disprove statements.

All four of the students used examples to evaluate the truth of statements and they all understood that constructing a counterexample is often the most efficient way to prove a statement is false. Although early in the semester the students often made mistakes recognizing and constructing counterexamples, they improved in this skill throughout the semester. However, even in those early days, they could express the sentiment that a counterexample is the only evidence needed to disprove a universally quantified statement.

Overall, the students could state the purposes of examples that Dr. S talked about repeatedly during the semester, namely using examples to understand statements and definitions, and using examples to evaluate the truth of statements. The students did not mention using examples to reveal logical inconsistencies or to make conjectures, and these purposes were modeled but not explicitly discussed as purposes of examples. However, the students did use examples for the purpose of making conjectures on the few tasks that did not provide the conclusion. The students seemed to know the purposes of examples discussed in class, and used examples for those purposes.

Summary of the purposes. The following purposes of examples were revealed through the data: 1) understanding a statement or definition, 2) making connections between definitions, 3) evaluating the truth of a statement, 4) making a conjecture, 5) generating a proof through a pattern or properties, and 6) revealing logical inconsistencies. The students primarily used examples for the purposes of understanding a statement or definition, and evaluating the truth of a statement. The students rarely used examples for the purposes of making a conjecture and generating a proof. The other purposes appeared only in the lectures.

Construction of Examples

In addition to knowing when and why examples might be useful, knowing how to accurately construct examples is of crucial importance. During this section, the construction of examples and counterexamples is discussed. Two levels of analysis were done on the students' responses: 1) were their examples actually examples, and 2) what construction techniques did they use. In addition, the instruction was analyzed for evidence discussing how to construct examples, both what characteristics does a construction need, and how does one go about actually constructing such an object. In these section the term construct indicates a mathematical object that does not satisfy the requisite characteristics to be an example or counterexample, or merely has not yet had those characteristics verified. The terms example and counterexample are reserved for constructions that have the necessary characteristics.

In terms of construction technique four categories were used. The first three are *trial and error*, *transformation* and *analysis* (Antonini, 2006). *Trial and error* refers to the process of selecting the parameters for the desired construction and then testing whether the needed characteristics hold. This category can subsume a wide variety of constructions. A prover who randomly selects integer parameters to test conditions uses *trial and error*, as does a prover who randomly picks the first parameter but then purposefully picks the remaining ones. The key characteristic is that in the *trial and error* technique, the prover does not know all the characteristics hold until after the construction is complete.

Transformation refers to the process of beginning with a construction that is known to satisfy some of the required characteristics, and then altering the construction so that all of the characteristics hold. The key feature is that a known example is altered. Often this technique occurs after an unsuccessful attempt at using *trial and error*. A prover uses *trial and error* to get an initial construction that has some of the conditions, and then *transforms* the construction into an example or counterexample.

Analysis is the process of analyzing the needed conditions of the problem to deduce the properties that an example must have. This type of construction is very rare, involves significant mathematical sophistication, and usually only appears after the previous two techniques have failed (Antonini, 2006). This type of construction was not observed from the students or the instruction through the entire semester.

The fourth category for this phase is not a technique for construction at all. An *authoritarian* example is retrieved from a source, instead of being constructed by the prover. For the students, the *authoritarian* sources included their lecture notes, their textbooks, and the task statements.

Amy. Amy was fairly proficient at constructing examples and counterexamples throughout the semester. She tended to be accurate with her examples, and fairly sophisticated in her technique.

Accuracy of construction. In general, Amy constructed examples that satisfied all necessary conditions of the problem. Only six of the 36 examples she generated were inaccurate; one example missed an hypothesis and five others were incomplete. An example was classified as incomplete when Amy was unable to construct a fully explicit example. For instance, Amy said "I know if [a function is]

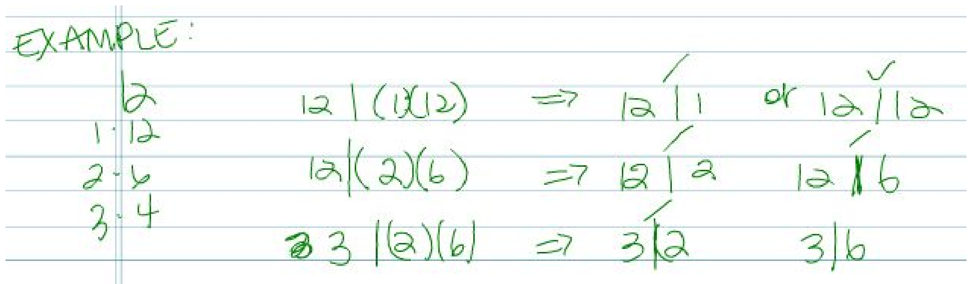

Figure 15: The accurate examples and counterexample constructed via *trial and error* by Amy in question 3 of interview 1.

periodic, then I can make it fine by twisting things." She knew that graphical transformations existed that would allow her to adjust a periodic function into a fine function, but she was unable to explicitly determine those transformations for a specific function in the interview setting.

> **Task** (Interview 1, Question 3) Provide either a proof or a counterexample for the following statement. For integers a, b, and c, if a divides bc, then either a divides b or a divides c.

Amy was the only student to accurately construct a counterexample on interview 1, question 3. As seen in Figure 15, Amy constructed two examples and a counterexample. When considering what value to pick for a, Amy mumbled "so... I ...want.... not a prime... I want a multiple of something... how about 12," where the ellipses indicate pauses not missing text. She appears to have selected the other values based on initially choosing 12. Amy did correctly assess the hypotheses and conclusions, although she did not label the example as examples or counterexamples until the end of the interview.

Construction technique. Amy utilized multiple techniques for constructing examples: three examples were from an *authoritarian* source, 17 were constructed by *trial and error*, and 16 were transformed from other examples. Amy generally had strong intuition for constructing examples. Even when using a *trial and error* technique, she often choose effective examples early in the process. For instance, when proving "if a divides bc, then either a divided b or a divides c" from question 3 of interview 1 (discussed above), Amy recognized that she did not want a prime number for the divisor, and immediately picked 12 as the divisor, see Figure 15.

> **Task** (Interview 2, Question 4) The number of four-digit numbers that can be formed using exactly the digits $1, 3, 3, 7$ is less than $4!$, because the two 3's are indistinguishable. Prove that the number of permutations of n objects, m of which are alike is $\frac{n!}{m!}$.

Amy frequently employed the *transformation* technique of example construction to generate examples. Two instances of this occurred during the second interview on question 4, which is in the field of combinatorics. Amy used several examples to explore question 4. She started with the *authoritarian* example $1, 3, 3, 7$ (given in the problem), and then transformed to similar examples, specifically $1, 3, 3, 3$ and $1, 1, 3, 3$. For each of these examples, Amy wrote the full list of outcomes and counted the number of permutations. Unfortunately, Amy was unable to connect these examples to the general statement in a meaningful way. During her attempts to generalize Amy tried to use combinations to develop the formula, not entirely remembering the correct formula or meaning of combinations. As a consequence, she was unable to shift from the specific examples to the general.

Carl. Although Carl did not use many examples, he was proficient at constructing those that he did use. He tended to utilize *trial and error* as his construction technique, but he always thought deeply about the choices he was making in the construction phase.

Accuracy of construction. Nearly all of Carl's constructions were accurate examples that satisfied all the required conditions. Only three examples contained inaccuracies; two examples were incomplete and one example construction did not satisfy the desired conclusion. One incomplete example occurred when Carl began constructing an example for a number theory question involving three integer values, but only selected a number for one of the three variables. The other incomplete example occurred when constructing a fine function. Carl drew a graph of a modified sine graph, but did not know how to modify the equation. This was classified as incomplete because Carl sought the algebraic form of the example, although the graph was a complete example.

> **Task** (Interview 3, Question 5) Let $f : \mathbb{R} \to \mathbb{R}$ be a fine function. If $g : \mathbb{R} \to \mathbb{R}$ is a function, such that $g \circ f$ is a fine function, what conditions must g satisfy? Does g have to be a fine function as well? What is the weakest condition that g must satisfy to ensure the composition is fine?

The example construction that did not satisfy the desired conclusion occurred on question 5 of the third interview. Carl reached the correct conclusion that $g(0) = 0$ without constructing an example. Then Carl decided to produce a "counterexample", specifically $g(x) = x + 1$ a function such that $g \circ f$ is not fine. This is not actually a counterexample to his claim; it is an instantiation of the contrapositive. This example does not satisfy the conclusion needed to be a counterexample, so it received the missing conclusion classification.

Construction technique. Carl usually used a *trial and error* technique for constructing examples. Two examples came from the statements of the questions, and two were constructed through a *transformation* process, but the remaining 15 were *trial and error* constructions.

> **Task** (Interview 1, Question 3) Provide either a proof or a counterexample for the following statement. For integers a, b, and c, if a divided bc, then either a divides b or a divides c.

On question 3 of the first interview, Carl began a counterexample construction by choosing $a = 4$. Since the statement is true for prime values of a, but not for composite values, Carl made an excellent choice for this parameter and the goal of finding a counterexample. Unfortunately, Carl chose $b = 16$ and $b = 12$ as the values for b. Since these are both multiples of 4, they satisfied the statement, and it did not matter what values were chosen for c. This was not productive to finding a counterexample. Carl did seem to think about the conclusion when selecting the values for his attempted counterexample. After generating these two examples, Carl concluded that it was not possible to find a counterexample, and attempted a proof.

Raul. Raul struggled with example construction during the first interview as he appeared to not know the properties required for examples and counterexamples. However, Raul seemed to improve in his skill of example construction throughout the semester.

Accuracy of construction. Most of the examples that Raul constructed were accurate. Of the 21 examples he created only three were inaccurate, and these all occurred during the first interview.

> **Task** (Interview 1, Question 2) Provide either a proof or a counterexample for the following statement. For integers a, b, c, and d, if a divides $b - c$ and a divides $c - d$, then a divides $b - d$.

In the first interview, Raul attempted to construct counterexamples for the *prove and disprove* questions. On question 2, Raul arbitrarily selected integer

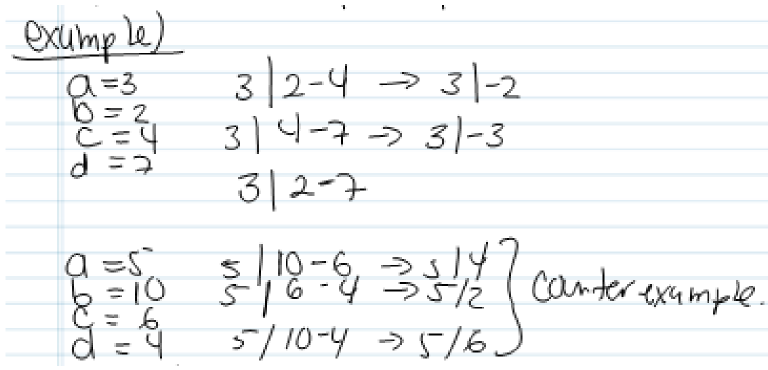

Figure 16: The examples generated by Raul in question 2 of interview 1.

numbers for the four variables. He did not consider the hypotheses or the conclusion until after the numbers were selected, although he did test if the assumptions held. Unfortunately, for the two sets of numbers that Raul selected, neither set satisfied both hypotheses. At the end of the first example Raul said "based on my example,the statement is not congruent." Raul did not clarify what he meant by not congruent, and he continued onto another example. Raul labeled the second set of numbers to be a counterexample. He argued that "the assumption is not correct, so this will be my counterexample" (see Figure 16). It appears that, at this point in the semester, Raul believed a counterexample only needed to violate the conclusion, and that the hypotheses did not need to be satisfied.

Construction technique. When constructing his examples, Raul used *trial and error* 11 times, *transformation* of examples eight times and took an example from an *authoritarian* source twice. The *authoritarian* source for both examples was the statement of the question.

Task (Interview 3, Question 3) Given an example of a fine function and explain why it is a fine function.

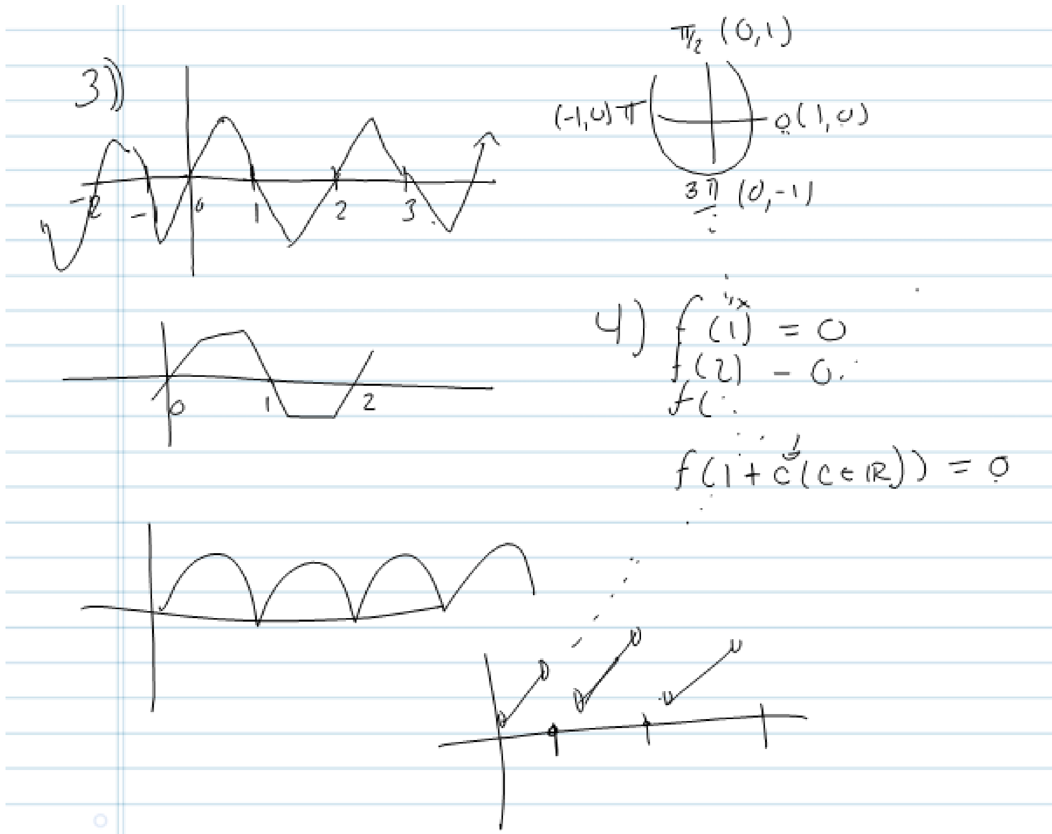

Figure 17: Raul drew several sketches of fine functions.

On the example generation of fine functions task, Raul repeatedly used the *transformation* of examples to complete this task. Raul immediately thought of transforming $y = \sin x$ to force the zeros to occur at the integers. He struggled to remember exactly how to transform the equation, so he drew a unit circle, and a sketch of the graph, (see the top of Figure 17).

Raul transformed this into several other graphs of fine functions. He started by drawing a "super weird" piecewise version of the previous function (see the middle left of Figure 17). Then he considered the absolute value of the first function (see the bottom left of Figure 17). Finally, he considered another piecewise function where the graph has discontinuities at the integers, in addition to zeros at the integers (see the bottom center of Figure 17). Raul intended for this graph to be non-periodic, but the graph does not show that feature particularly well.

Mike. Mike struggled with constructing useful examples. Unfortunately, since he did not complete the third interview, and used only one example in the second interview there is no evidence to say whether Mike showed improvement throughout the semester. When Mike used *trial and error*, he tended to think very little about the values he was picking for the parameters.

Accuracy of construction. Four of Mike's examples were accurate. For the remaining two examples, one example missed the hypotheses and the remaining example was incomplete.

> **Task** (Interview 2, Question 1e) Claim. If the relation R is symmetric and transitive, it is also reflexive.
> "Proof:" Since R is symmetric, if $(x, y) \in R$, then $(y, x) \in R$. Thus $(x, y) \in R$ and $(y, x) \in R$, and since R is transitive, $(x, x) \in R$.
> Therefore R is reflexive.

The incomplete example occurred on the argument evaluation task shown above. In an attempt to understand the argument, Mike constructed an alternate example using the set $\{a, b, c, d\}$ as the domain and codomain for a relation R (see Figure 18). Mike mixed up the notation for relations using the notation $(a, b)R(b, a)$ when he intended $(a, b) \in R$ and $(b, a) \in R$ or aRb and bRa. Mike used the notation for $R : A \times A \to A \times A$, but talked about it like the relation $R : A \to A$.

Mike began his construction by considering the hypotheses of the relation being symmetric and transitive. The line $(a, b)R(b, a)$ is considering the symmetric property, and the lines $(a, b)R(b, c)$ and (a, c) were for the transitive property. However, even when the notational difficulties are ignored, Mike's example did not

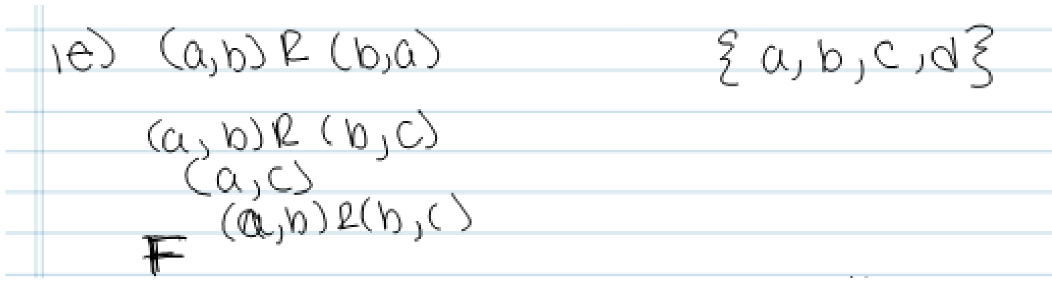

Figure 18: Mike struggled with the notation of relations, and did not construct the entire relation.

include all of the ordered pairs necessary to be symmetric or transitive. Mike abandoned the example, concluding "this one fails,... the reason it fails is because you can provide a counterexample, where the claim doesn't work." He knew that he had not actually constructed such a counterexample, but he did not say why he knew one existed.

Construction technique. Mike used *trial and error* for four of the examples and a *transformation* technique for two examples.

Although Mike was fairly careful about checking the hypotheses in his *trial and error* examples, he was not as careful about what he could learn from his examples. For instance, in the task displayed below Mike decided to seek a counterexample. He selected the value $a = 1$ for the divisor "because one divides everything". Mike did not realize that the selection will prevent the formation of a counterexample.

Task (Interview 1, Question 2) Prove or disprove with a counterexample. If $a|(b-c)$ and $a|(c-d)$, then $a|(b-d)$.

After Mike realized that the numbers he chose formed an example, not a counterexample, and he attempted a proof. Mike wrote a correct proof of the statement, so he did not attempt to construct another example. Mike never realized the contradiction between his desired goal of finding a counterexample, and the value he chose for a.

Comparison of the students. The construction of examples was a difficult task for many of the students. During the first interview, Raul and Mike both made errors in constructing examples because they did not know which conditions a construction needed to satisfy to be classified as an example or a counterexample. In fact, they both identified constructions as counterexamples that did not satisfy the hypotheses of the statement. However, the students showed improvement in these skills later in the semester.

Although the students generally constructed examples that were accurate in a logical sense, the often did not construct useful examples. Additionally, they often did not recognize why their examples were or were not useful. This behavior was particularly prevalent in the first interview, on questions 2 and 3, which are included in the indications section of the cross-analysis. Amy was the only student who answered both of these questions correctly. All three of the others had examples of mixed usefulness.

On question 2, Mike was seeking a potential counterexample and chose $a = 1$ as the value for the divisor, stating that he chose this value because "1 divides everything." Mike did not realize that this choice for a meant that every possible example would be true. Luckily for Mike, this particular claim is true and he wrote a complete proof of the claim. However, it is troubling that Mike did not consider the implications of his choices. This was also the only instance that a student constructed an example in such a way that directly contradicted their intended purpose. Although other students constructed examples that were not useful for their purpose, this was the only instance in which a student stated a fact that would directly indicate the lack of usefulness, yet did not recognize the contradiction.

Carl experienced a similar issue on question 3. Carl was seeking a counterexample of the statement, and generated the following constructions $a = 4, b = 16, c = 4$ and $a = 4, b = 12$. In both circumstances, Carl picked a value for b such that $a|b$. Since the conclusion of the statement says $a|b$ or $a|c$, Carl's choices for these two values automatically satisfy the conclusion. Carl even realized that the second example satisfied the conclusion before he selected a value for c, but he did not realize that his choices for b prevented him from constructing a counterexample.

Table 4: This table summarizes the construction abilities of the students.

Construction	Amy	Carl	Raul	Mike	Total
Accurate Construction	30	16	18	4	68
Inaccurate or Incomplete	6	3	3	2	14
Authoritarian Source	3	2	2	0	7
Trial and Error	17	15	11	4	47
Transformation	16	2	8	2	28

In addition to improving the accuracy of their constructions, the students began to exhibit more advanced construction techniques later in the semester. During the first interview, Mike was the only students to utilize the *transformation* construction technique, and he only did so once. By the final interview, the students were using the *transformation* technique more frequently than *trial and error*. This interpretation of the result may be conflated with the choices the students make due to the mathematical content. Specifically, the first interview consisted entirely of number theory tasks which the students may have limited previous experience, whereas the final interview concerned real-valued functions and the students should have significant experience with these from their calculus courses. Although the students likely used the *transformation* technique due to increased experience, they also knew more examples of real-valued functions to draw upon as the starting point for the transformation process.

In particular, when asked to construct an example of a fine function on question 3 of interview 3, the first example constructed by each student was a transformation of $y = \sin x$. These students recognized that the pattern of the zeros in $y = \sin x$ could be adjusted to satisfy the conditions of a fine function. It is unlikely that the students could have constructed an example of a fine function via *trial and error* because of how difficult it would be to verify. However, it is equally difficult to imagine a students utilizing a *transformation* technique on $a|(bc)$ implied $a|b$ or $a|c$, especially for an initial example of the statement. Most students will not have a sufficient background in the formal language of divisibility to have such examples in their personal example space.

Dr. S's modeling. Dr. S modeled example construction very rarely during the lecture. Although she presented many examples throughout the semester, she seldom talked about how these examples were constructed. Dr. S did model how to determine which properties an example or counterexample of a statement needs to satisfy, and how to go about verifying that a construction satisfies those properties. Dr. S knew that *trial and error* is the first technique used in example construction, and that the most important aspect of that is knowing which properties need to be verified. Dr. S intended for the board work problems that she assigned to be opportunities for the students to learn how to construct examples. She knew that the students would often fail before they succeeded at example construction, and that the best way to help the students improve would be to review their constructions attempts during class.

There were two episodes from the lecture where Dr. S emphasized example construction, and the care that must take place when constructing examples. The first instance occurred shortly after defining functions. Dr. S emphasized the importance of a function being well-defined, particularly when the domain is a partition. To do this, Dr. S presented three potential functions:

$$\begin{aligned} &f:\mathbb{Z}_3\to\mathbb{Z}_6 && f\left([x]_3\right)=[3x+2]_6\\ &g:\mathbb{Z}_4\to\mathbb{Z}_2 && g\left([x]_4\right)=[3x]_2\\ &h:\mathbb{Q}\to\mathbb{Z} && h\left(\tfrac{a}{b}\right)=a+b \end{aligned}$$

The first example was generated using numbers suggested by the students, the last two were purposely chosen by Dr. S. Dr. S showed that f and h are not well-defined by producing counterexamples that show that two different representatives of the equivalence classes produce different outputs, specifically $f([0]_3)=[2]_6$ but $f([3]_3)=[5]_6$, and $h\left(\frac{2}{3}\right)=5$ but $h\left(\frac{4}{6}\right)=10$. For g, Dr. S provided the students with a proof that it was well-defined. Ultimately this episode was demonstrating what it means to be well-defined, but Dr. S knew that this would help the students when constructing their own functions especially in their Modern Algebra course. In fact, Dr. S explicitly told the students that she was emphasizing well-defined here, because she often wants the students to know this topic better when she teaches Modern Algebra.

The other episode of example construction occurred late in the semester. After introducing the definitions of one-to-one and onto, Dr. S led the class through a demonstration of the *transformation* technique. Dr. S instructed the class to alter the domain and codomain of the function $f(x)=x^2$ to make it one-to-one and onto. A transcript of the episode follows.

> Dr. S: An interesting answer to your question about one-to-one and onto is you can take one function and change the domain and the codomain to determine whether it is one-to-one or onto. ... Let's suppose we took $f:\mathbb{R}\to\mathbb{R}, f(x)=x^2$. If I asked the question is that one-to-one, what's your general answer? Why not?

Dr. S: All right, so it's not one-to-one because one... Somebody actually said it so I should write it down. So $f(1)$ equals $f(-1)$ but $1 \neq -1$ so this is not one-to-one if $f : \mathbb{R} \to \mathbb{R}$, right?

Richard: If you were to say $f(x) = f(y)$ and you put $x^2 = y^2$... Oh, when you take the square root it is ± 1, okay.

Dr. S: ... Hey, Richard what did you just say?... He said suppose $f(x) = f(y)$ then $x^2 = y^2$. In the naturals I can just take the square root and everything's positive so I can get $x = 1$. So he's proven it. What about onto? Over here is $f : \mathbb{R} \to \mathbb{R}$ onto? Prove it. ... Let $w = -1$. If $f(x) = w$, what is x?

Student: It's not defined.

Dr. S: It's not in $\mathbb{R}$. It is actually complex. It's not in your domain. There is no element of your domain that maps to -1. So this is not onto. ... All right, so over here going to the naturals fixed things for us. Is [$f : \mathbb{N} \to \mathbb{N}$] onto now? Wait a minute, you're saying no. How come?

Student: Because the square root of one is not...

Dr. S: But negative one's not [in $\mathbb{N}$]. But keep thinking on those lines, keep going. [inaudible] Well somebody just said two or three. [inaudible] Let $w = 2$. If $f(x) = w$, x is not a natural number. What would x be if $f(x) = 2$, what's x?

Student: Irrational.

Dr. S: The square root of two, which is irrational. So it's certainly not a natural number. There is no natural number that when you square it you get two. So [$f : \mathbb{R} \to \mathbb{R}$] is not onto. [$f : \mathbb{R} \to \mathbb{R}$] is not one-to-one here. [$f : \mathbb{N} \to \mathbb{N}$] Is one-to-one there. Not onto. Here is a one-to-one not onto.

Dr. S: I want you to fill in the blanks, f : _____ $\to$ _____, so that it is not one-to-one and is onto. Same x. $f(x) = x^2$. I want you to change your domain in front of me. Now I want it not to be one-to-one but I do want it to be onto.

Student: Natural numbers

Dr. S: You're thinking on the right lines you just haven't gotten it yet. [inaudible] The reals, we know that one-to-one is going to fail if I start with the reals. So that's a good place to start. So over here if I start with the reals there's my proof that you're not one-to-one. Now I want to be onto. ... Student: The irrationals?

Dr. S: Irrationals? Negative pi. Integers. How do I get negative one? That's an integer. [inaudible] Reals? That's what we tried here. What messed up onto?

Student: Negative one.

Dr. S: Negative one. What else might mess up onto? Anything negative. So how do I make this thing onto?

Student: $\mathbb{R}^+$

Dr. S: Almost positives. Almost $\mathbb{R}^+$.

Amy: $\mathbb{R}^+$ plus 0.

Dr. S: Yes. $\mathbb{R}^+ \cup \{0\}$. So I need 0 in there. $\mathbb{R}^+$ is the positive real numbers. So the output of x^2 is all real numbers, greater than or equal to zero. This one is not one-to-one because the negatives, one and negative one both go to the same place and it is onto. All right, one more. Fill in the blanks, $f : \_\_\_\_\_ \to \_\_\_\_\_$. I want to fill in the blanks so it's one-to-one and onto.

Student: Natural numbers to...

Student: Naturals to reals.

Dr. S: If I start with the naturals and I end with the reals I'm good with one-to-one. But just like over here when I end with the reals, 2 is a real. I would have to end with only those numbers that are perfect squares. That's really nasty. Let's start with something bigger than natural numbers but not as big as all the reals.

Student: Integers.

Dr. S: You're still only going to get squares. But you could do it that way but let's start with something like... Actually if I just go $\mathbb{R}^+ \to \mathbb{R}^+$. Let's throw away zero today. Let's go $\mathbb{R}^+ \to \mathbb{R}^+$. If I only start with positive numbers, I don't even count zero then I'm one-to-one. Yeah. Am I onto all positive real numbers?

Student: Yes.

Dr. S: I can take a square root of any positive number and get a real, which is back in my world. Now have I more than thoroughly answered your question, can a function be one-to-one or onto or both or neither?

Richard: Yeah, absolutely.

This vignette shows Dr. S guiding the class through the *transformation* technique of example construction, as they are taking a single example and adjusting is slightly to satisfy the needs of the current problem. Of particular note is the large number of incorrect guesses given by the students. As Dr. S commented in her interviews, it was her opinion that some of the students are prepared for *transformation* and *analysis* techniques, but that some of the students were not ready to move beyond *trial and error*. This belief seems to be supported by the episode, since a handful of students seemed to be shouting out random suggestions rather than carefully considering the problem.

Dr. S's discussion. Throughout the course, Dr. S seldom lectured explicitly about constructing examples and counterexamples. One reason that Dr. S did not talk about example construction in the lecture is because many of the exercises assigned for board work concerned example construction. Dr. S had the expectation that the students would attempt and present those questions on the board, and that these presentations would provide the opportunity to discuss example construction for the various definitions. The frequency of the board work presentations varied throughout the semester, with many presentation during the first few weeks and after the first test, and fewer presentations in the weeks before the first test and after the second test. When the students complained that they did not have a sufficient number of examples, Dr. S replied "do more board work questions." However, a consequence of the students not presenting enough board work problems is that the class missed opportunities to talk about example construction.

Another reason that Dr. S did not lectured about example construction frequently is because she expected the students to utilize *trial and error* by randomly trying constructions and to test whether these are examples or not. Although this is not a sophisticated strategy for example construction, Dr. S believed that students at the earliest stages of proof writing "are not always ready yet" for other strategies. Dr. S wished that the students would move towards the *transformation* construction strategy by asking themselves questions such as "is the statement similar to one [I] know?" and then using that response to construct their example. During the final interview, Dr. S reiterated this by saying "I would like to move them toward more directed examples where they are intentionally trying to go certain places but I doubt that most of them are ready for that. Right now I'm happy if they try random examples to see what's going on, as long as they don't stop there." Perseverance was a frequent theme when discussing proof and example constructions in the lecture.

Most of the discussion of example construction occurred during the review of student presentations, rather than during the lecture components. When Dr. S reviewed the student presentations, she would ask questions of the students for clarification and to expand their thinking. For instance, one student presented a task that asked her to describe four different partitions for the set of university students. The student divided the set of students into the set of freshmen, the set of sophomores and so on. She then presented four partitions using these subsets, e.g. $\{\{\text{freshmen}\}, \{\text{sophomores}\}, \{\text{juniors}, \text{seniors}\}\}$ and $\{\{\text{freshmen}, \text{sophomores}\}, \{\text{juniors}, \text{seniors}\}\}$. Dr. S responded to these asking the entire class to come up with other partitions for university students. The students responded with a variety of responses, including major, number of hours completed and grade point averages. Dr. S used questions to get the class to consider the question more fully.

When reviewing student work, Dr. S frequently asked for clarification from the student presenters. One time that Dr. S did this occurred early in the semester, when a student answered the task below with the statement "If I can beat Lebron James in a 1-on-1 game of basketball, then I can beat my friend Carlos 1-on-1."

> **Task** Give, if possible, an example of a true conditional sentence for which the converse if false.

After a quick chuckle, Dr. S claimed that she could not asses whether this example satisfied the desired characteristics of the task without additional information. Dr. S turned this interchange into a brief discussion of the importance of careful writing and provided sufficient details in proofs. This is an amusing anecdote, but Dr. S often asked other questions of the students of a more serious nature. When a student presented a correct example or counterexample, Dr. S would ask how the student came up with their examples in order to assist the rest of the class.

When the students presented example construction tasks that we incorrect, Dr. S would usually ask the student who presented (or sometimes the whole class) to help her revise the construction. In one instance of this phenomenon Carl presented a relation on $A = 1, 2, 3$ that should have the properties of symmetry, transitivity and not reflexivity. Carl presented the relation $\{(1,2)(2,1)(1,3)(2,3)(3,2)(3,1)\}$, but this example is not transitive. Dr. S argued that if $(1,2)$ and $(2,1)$ are in the relation, then transitivity requires that $(1,1)$ and $(2,2)$ must also be included. As such, Dr. S changed the relation to $\{(1,2),(2,1),(1,1),(2,2)\}$, which is symmetric and transitive, but not reflexive because it is missing $(3,3)$. Although Dr. S did this with the intention of correcting Carl's work, she essentially walked the students through using the *transformation* technique for example construction, since she transformed an existing example to satisfy the given criteria.

In summary, Dr. S rarely addressed example construction when lecturing, preferring to address the constructions during the student presentations. Additionally, Dr. S believed that most of the students in the course were only prepared to attempt example construction with a *trial and error* technique. As a result, when she did talk about example construction it was usually to provide the students with an opportunity to experience the *transformation* technique. Although, Dr. S provided a considerable number of examples to the class, she seldom explicitly discussed how she came up with such an example.

Comparing the instruction and the students. Two techniques for example construction were used by both the instructor and the students, *trial and error*, and *transformation*. The *analysis* technique was not demonstrated by the instructor or used by the students; however, during the member checking interview, Dr. S argued that the *analysis* technique was too advanced to be useful to the students at their current level of understanding.

The students used the *trial and error* construction technique for all of the examples constructed during the first interview, with one exception. However, as the semester progressed the students used the *transformation* technique with increasing frequency. Dr. S predicted this behavior of the students. In the first interview, Dr. S said

> It depends on the problem, but to some extent, *trial and error* is the very first step. You just try stuff. I've seen this even with advanced REU [Research Experiences for Undergraduates] students, where there is a good strategy. They're not always ready yet. I'm okay with them randomly trying at first. Now, I want them to move toward more careful construction. As they go through this, they should be looking for things that are similar and using that to give them a hint.

Dr. S recognized that as beginning students, they would not have the mathematical experience to use the more advanced *transformation* and *analysis* techniques, but she hoped they would grow to that point. During the same interview, Dr. S elaborated that although she expects the students to have some familiarity with using examples from their calculus classes, "they just never had to construct [examples] themselves before." As such, some of the difficulties the students had with example construction were expected.

During the final interview, Dr. S elaborated further on her expectations regarding the techniques used by the students for constructing examples.

> Realistically, I expect them to start random where they just pick some numbers, shove them in and see what happens and if they get what they want or what they need, they're done and if not, they pick some others. A little more idealistically, let's start random and get information when

> the random doesn't work on how to pick something that does. Okay, I picked a random example and I worked it through and it did work so it's not a counterexample but I see here I need to divide, let me put a zero in my example that won't let me divide, aha, I found my counterexample. Use information from the example to generate the next choice and that specific one was if they're doing counterexamples, but even if they're doing examples for understanding, random can start and then there's some piece of it you're not understanding and not seeing, why do I need this hypothesis? None of my examples need it; all of my examples have positive numbers let me try some negative ones. Seeing general patterns.
>
> Ideally, I want them to move toward and I doubt even the best ones are going to really maybe one or two very top students will get there by the end of the semester but move toward looking at a statement and being able to choose a non-random example to start. If I want the composition of functions to be one-to-one with the functions themselves not being one-to-one, how do I start generating an example? Thinking about the theory behind when you get one to one and what has to happen and what can't happen. ... I would like them to start choosing more narrow examples that fit the constraints of the problem. Honestly, I think right now the very best students may be pretty close to that. The average students might do one or two generic and then try that out and the weaker students are just in sheer generic.

This quote shows that Dr. S expected the students to grow to the more sophisticated techniques of example construction as they proceeded through this course, and into their subsequent courses. This is consistent with the work of the students in the interviews, and when Dr. S demonstrated the techniques during the semester.

Dr. S did not vocalize an expectation of the accuracy problems exhibited by some the students during the initial interviews. Both Raul and Mike had created examples that violated the statement hypotheses. Raul did not seem to realize that failing the hypotheses was a problem. Additionally, Mike and Carl struggled with choosing *useful* examples, meaning examples that improved their progress through the task. During the member checking interview, Dr. S said students often make these types of construction errors at this point in their development. She furthered this by explaining that many students present counterexamples that are not actually counterexamples, especially on the first test of the course.

Dr. S usually did not talk about her construction techniques when she presented examples to the class. This was by design with the intention for students to attempt to find examples via *trial and error*, and then to focus on correcting them if needed. Dr. S expected the students as individuals to solve each board work questions assigned, and as a class to present most of the board work problems assigned. She designed the course so that most of the example generation and counterexample construction tasked were assigned as board work, and that she would talk about example construction as she reviewed and corrected the examples in the presentations. However, many days the class included fewer than four student presentations and the students tended to present problems asking for proofs rather than the problems asking for examples. Consequently, Dr. S did not have the opportunity to talk about construction techniques with the expected frequency.

Towards the end of the semester, Dr. S showed the students a *transformation* technique on a limited number of examples, as discussed in more detail in the earlier section on examples in the course. She did this partly because the particular definitions and examples lent themselves to the technique, but she also to introduce

Table 5: The construction techniques used by the students during the interviews

Interview	1	2	3
Authoritarian Source	0	5	2
Trial and Error	23	6	18
Transformation	1	5	22

the construction technique to the students.

During the first interview, the students used *trial and error* almost exclusively. The second interview was nearly equal between *trial and error*, *transformation* and taking an example from an *authoritarian* source. In the third interview the students used a *transformation* technique more frequently than *trial and error*, see Table 5. It is clear that the sampled students became significantly more comfortable with this technique, but it is unclear why this occurred. It is likely several factors influenced this change including: the instruction provided in using a *transformation* technique, the increased amount of experience with constructing examples and proofs, and the change in the content. During the first interview, the content was all number theory questions where examples consisted of selecting integers that satisfied the given hypotheses. During the third interview, the content was real-valued functions, where examples could be represented by graphs, equations or tables. In addition, the students were familiar with a large family of real-valued functions due to their successful completion of the algebra to calculus sequence. It is likely that the variety of representations and the transformations possible on functions influenced this result.

Overall, Dr. S had the experience to know the capabilities of the students with respect to example construction. She recognized that *trial and error* would be the primary technique at the beginning of the semester, and that many of the students would not be able to move beyond that technique in this course. However, towards the end of the semester, she introduced the *transformation* construction technique for the benefit of the students who were ready for more advanced techniques. The students in the sample were able to apply the *transformation* technique in some circumstances, and likely will be able to utilize it more frequently in their subsequent courses.

Summary of the example construction techniques. The students in this course primarily used *trial and error* to construct their examples at the beginning of the semester. Sometimes the students appeared to blindly select integers for their constructions and then test the hypotheses and conclusions. Other times the students put more thought into their choices before testing the necessary properties. Eventually, many of the students began utilizing *transformation* techniques to alter other constructions. These techniques matched Dr. S's expectations for her students. She expected that they would start the semester by testing random examples, but that they would become more purposeful in their examples over the course of the semester.

Implications of Examples

The implications are the conclusions or decisions drawn by the prover after an example is constructed. The implications are often similar to the purpose of the example. In this study, the purpose is established before an example is constructed, but the implication is the reaction after the construction. A common implication is to fulfill the purpose, but sometimes a prover draws conclusions other than the stated purpose.

The implications from using examples can have many forms. The conclusions made are often related to the quantifiers in the statement. In particular, the quantifiers of the statement help to determine whether the implications are logically supported or not. Some common implications are making a decision about the truth

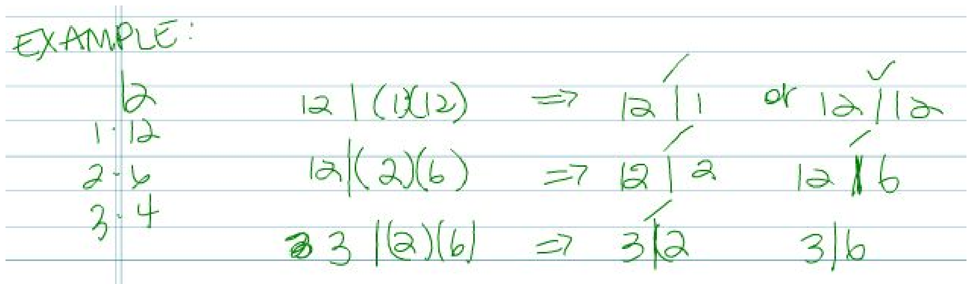

Figure 19: Although Amy constructed a counterexample on question 3 of interview 1, she did not recognize that the construction proved the statement false.

of a statement, making a conjecture, writing a proof, and doing nothing.

Amy. After constructing an example, Amy's most common reaction was to create another example, which she did 15 times. The other reactions were to decide the truthfulness of the claim (8), to connect to formal proof language (4), to make a conjecture (1), to reach a false conclusion (2) and to move on without drawing any conclusion (4). In two instances, Amy had constructed an example, but the interview was ended before she reached a conclusion.

> **Task** (Interview 1, Question 3) Provide either a proof or a counterexample for the following statement. For integers a, b, and c, if a divides bc, then either a divides b or a divides c.

Amy sometimes struggled with determining implications of the examples she used. The first time this occurred was during the first interview, while working on question 3. After reading the question, Amy generated two examples and one counterexample to the statement (see Figure 19). Unfortunately, Amy commented "I feel like it works, kind of" and then attempted to prove the statement by contradiction with an incorrect set up for the proof. As a result, Amy thought she had proven the statement.

At the end of the first interview, I asked Amy to review her work on this question. The following excerpt comes from this reflection.

> SH: Okay. The question said, "Prove or disprove." Right now what do you think you've done?
>
> Amy: Disproven it? Well, I think I've done both!
>
> SH: Okay.
>
> Amy: I mean ... I feel like this part is correct and the sense that I can get a times something to equal bc but I get a^2 in here and I don't think I'm allowed to have a in terms of a and things like that but I've already ... I guess I came up with an example, multiple examples here where it doesn't work.
>
> SH: You came up with two examples where it worked and one example where it doesn't.
>
> Amy: Yeah, like this one it will never hold but this oh yeah it says it works. And that works. I don't know. I don't know.
>
> SH: How do you disprove something?
>
> Amy: Come up with one example where it doesn't work. Done.
>
> Amy: Is that really all I needed? I say one ...
>
> SH: That's really all you needed.
>
> Amy: Cool.

When asked directly how to disprove a statement, Amy quickly responded with the correct strategy of generating a counterexample. However, during the problem Amy was unable to recognize that she had generated a counterexample and continued to work on the problem. This indicates a consistent behavior for Amy; she frequently demonstrated keen instincts for useful examples, but then struggled with recognizing what she had generated.

A few times Amy generalized too much from limited examples, especially working with fine functions in the final interview. In particular, Amy mistakenly assumed that all fine functions are periodic (question 4a, Appendix F). She reached this conclusion after considering only the sine function, and did not prove it in general. This error was compounded by using the statement in the proof of question 4c, which asks if a horizontal translation of an integer magnitude of a fine function will produce a fine function. Amy's proof of the horizontal translation question contained logic errors with quantifiers, and the false premise that all fine functions are periodic. After some prompting, Amy attempted to remedy the errors with the quantifiers and she realized that not all fine functions are periodic and drew a counterexample. Although Amy eventually considered several examples of fine functions, her initial trouble generating examples caused her to assume statements without actually considering a general argument.

Amy drew several accurate conclusions from the examples she constructed. For instance, Amy accurately constructed counterexamples to several statements. She also was able to use examples to clarify her understanding of definitions. For instance, after drawing the sketches shown in Figure 20 Amy realized that the formal definitions of increasing and decreasing function matched her intuition. With that information, Amy was able to construct counterexamples for both part of question 1 of the third interview.

Although Amy made some successful implications, she often did not understand what she had accomplished from her examples. Her default behavior was to create another construction, which may be a consequence of not knowing the implications of her work. Determining the implications was Amy's weakest phase. While she could state appropriate purposes of the examples, she often forgot these purposes when she had an actual example or counterexample in front of her.

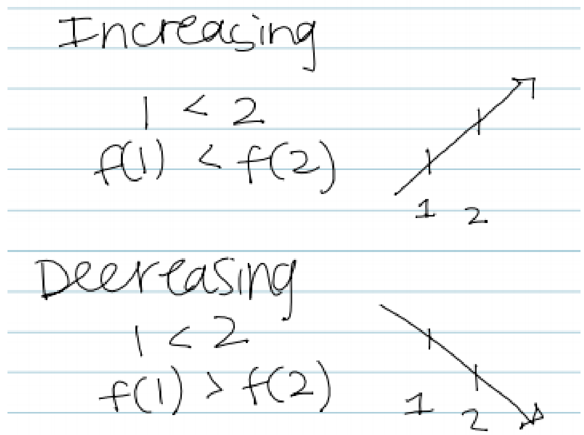

Figure 20: The examples drawn by Amy to understand the definitions of increasing and decreasing during question 3 of interview 3. From these sketches, Amy was able to accurately apply these definitions.

Carl. Carl made a variety of implications after constructing an example. In six instances, he constructed an additional example to gain insight, although all of these occurred in the first interview, and five were on a single problem. After six other examples, Carl evaluated the truthfulness of a statement and either concluded that one of the examples is a counterexample or then attempted a proof. Carl used two example constructions to connect to the formal proof language needed to write a proof. Additionally, in four instances Carl did not draw any conclusions, and in one instance he drew a false conclusion.

> **Task** (Interview 1, Question 2) Provide either a proof or a counterexample for the following statement. For integers a, b, c, and d, if a divides $b - c$ and a divides $c - d$, then a divides $b - d$.

During the first interview, Carl constructed six examples while working on question 2. Carl arbitrarily selected values for a, b and c that satisfy the first hypothesis. Then he tried to determine a value for d that would form a counterexample (see Figure 21). When Carl talked about the numbers he chose for

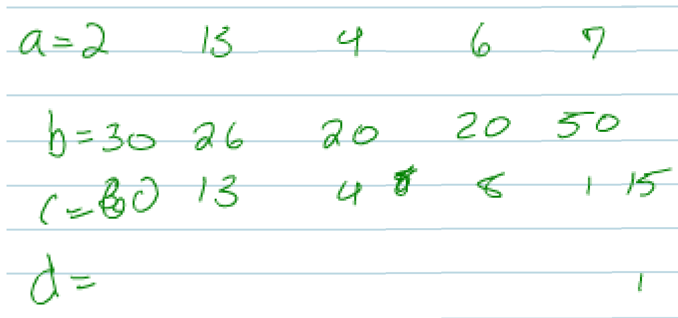

Figure 21: The list of examples generated by Carl in question 2 of interview 1.

a he said "they were arbitrary," but he "knew if it got too big, it'd be too difficult to come up with solutions for b, c, and d." Carl decided that the statement was true, although he was unable to produce a satisfactory proof for this task. Just after the interview had officially ended, Carl asked about the proof of this statement. When I indicated to him that the main idea is simply that $(b-c)+(c-d)=b-d$, Carl replied that he saw that pattern when trying to generate a counterexample, but that he "didn't know you could do that." Carl was unable to translate the pattern he observed into the language necessary for a written proof.

Task (Interview 2, Question 3) Let $\mathcal{A}=\{A_\alpha : \alpha\in\Delta\}$ be a family of sets, $\Delta\neq\emptyset$, and B be a set. For each statement either prove the statement is true or give a counterexample.

(a)
$$B-\left(\bigcap_{\alpha\in\Delta} A_\alpha\right)=\bigcap_{\alpha\in\Delta}(B-A_\alpha)$$

(b)
$$\left(\bigcup_{\alpha\in\Delta} A_\alpha\right)-B=\bigcup_{\alpha\in\Delta}(A_\alpha-B)$$

During the second interview, Carl used an example to evaluate the truth of both parts of question 3. Carl began by attempting to generate counterexamples. On part a, Carl generated the family of sets $A_1 = \{1, 4\}, A_2 = \{1, 3\}, A_3 = \{1, 2\}$ and $B = \{1, 2, 3\}$. After performing the set operations for each side of the statement, Carl verified that these sets are a counterexample to statement 3a. Carl then used the same sets for part b. However in this circumstance these sets resulted in an example. Then Carl adjusted set $B = \{1\}$, which produced another example. At this point, Carl decided that the statement might be true and wrote an accurate proof of the statement. Carl did not explicitly refer to the example at any point during the process of writing the proof, but it appears that by this point later in the semester Carl had improved his ability to connect examples and formal proof.

This improvement extended even further in the third interview, when Carl accurately connected the example $(\sin x) \cdot x^2$, to the proof that the product of a fine function and any other function is fine. In this instance, Carl directly referred to his example while writing his proof, and the end results was an accurate proof. These three instances viewed together show that Carl improved in connecting examples to formal proof language throughout the semester.

Raul. As a consequence of an example, Raul constructed an additional example six times, he decided the truthfulness of the statement six times, and he connected the example to formal proof language three times. However, Raul also did not draw any conclusion from his examples five times and made a false conjecture once. It appears that Raul knew how examples could help him on tasks, yet he struggled with exactly which conclusions he could draw.

> **Task** (Interview 1, Question 1) Let a, b, and c be natural numbers and $\gcd(a, b) = d$. Prove that a divides b if and only if $d = a$.

Raul was the only student who was able to connect an example to formal proof language during the first interview. When Raul was working on question 1, he constructed the example $a = 2$ and $b = 4$, so that $d = \gcd(2, 4) = 2$. After constructing and thinking about this example, Raul realized that a consequence of $a = d$ is that a is also the greatest common divisor of a and b. From this observation, Raul was able to prove the "if $a = d$, then a divides b" direction of the biconditional statement. Although, Raul was unable to complete the other direction of this proof, the insight he gained from considering this example allowed him to make more progress than he would have without the example.

In the final interview, Raul drew several examples of fine functions in order to gain more insight into the definition, see Figure 17. However, Raul did not use these examples to help him solve the tasks in the subsequent questions. In fact, Raul did not use any examples in the remaining questions, and ultimately drew no implications from these constructions.

Mike. Mike created another example twice, decided whether a statement was true or false after three examples, and did not draw a conclusion in the final instance. He never connected an example to a formal proof.

> **Task** (Interview 1, Question 3) Provide either a proof or a counterexample for the following statement. For integers a, b, and c, if a divided bc, then either a divides b or a divides c.

Although Mike did not ever connect an example to a formal proof, he did recognize that examples can inform proof writing. After reading the task shown above, Mike decided to begin with a proof attempt rather than an example. However, after he got stuck in the proof, Mike decided to try for a counterexample. He chose the value $a = 2$, and then $b = 2$ and $c = 3$. Mike realized that this selection was an example of the statement, but not that his selection of a prime for the divisor a prevented him from constructing a counterexample. Mike then returned to proof writing. Mike believed he constructed a correct proof, however, he proved the converse of the statement rather than the statement itself.

Comparison of the students. As seen in Table 6, the students often attempted to construct multiple examples and counterexamples when they chose this behavior. The most frequent behavior after constructing one example is to construct another example. Other frequent behaviors include deciding the truthfulness of the claim (sometimes correctly and sometimes not), and doing nothing (i.e., drawing no conclusions and either ending the interview or moving onto the next question). This component of using examples seemed to be the most difficult for the students, as they often struggled in this area. Arguably, the two most valuable implications for proof writing are evaluating the truthfulness of a statement and connecting examples to the formal proof. As such, these two implications will be discussed in greater detail below.

Table 6: This table summarizes the implications exhibited by the students.

Implications	Amy	Carl	Raul	Mike	Total
Another Example	15	6	6	2	28
Decide Truthfulness	8	6	6	3	23
Made a Conjecture	3	1	1	0	6
Connect to Proof	4	2	3	0	9
Did Nothing	6	4	5	1	16

Evaluating the truthfulness of a statement. The students collectively had mixed results in evaluating whether statements were true or false. As seen in Table 7, each student reached an erroneous conclusion at least once in the course of the interviews. In the table, the phrase *Believed True* indicates the student believed the statement to be true but knew their proof was incomplete, *"Proved" True* indicates the student believed they had proved the statement true but the proof was incorrect, and *Proved True* indicates the student successfully proved the statement. At other times the students were unable to articulate whether or not they believed the statement to be true or false.

Table 7: This table lists the conclusion about truthfulness reached by the students on the *prove or disprove* tasks, and whether they constructed examples to reach the conclusion. The first column lists interview number, question number and the actual truthfulness. A * indicates that the student believed the construction was an example or counterexample, but it actually was not.

	Amy	Carl	Raul	Mike
1.2 (True)	Proved True None	Believed True Examples	False Counterexample*	Proved True Example
1.3 (False)	Unsure Counterexample	"Proved" True Examples	Believed True Example	"Proved" True Example
2.3a (False)	False Counterexample	N/A	False Counterexample	N/A
2.3b (True)	Unsure Example	N/A	Proved True Example	N/A
3.1a (False)	False Counterexample	False Counterexample	False Counterexample	N/A
3.1b (False)	False Counterexample	False Counterexample	Believed True Incomplete Proof	N/A
3.2 (False)	False Counterexample*	Unsure None	False Counterexample	N/A
3.4a (False)	"Proved" True Example	False Counterexample	Unsure None	N/A
3.4b (True)	Proved True Example	Proved True Example	Proved True None	N/A
3.4c (True)	Believed True Incorrect Proof	Believed True Incorrect Proof	N/A	N/A

Task (Interview 1, Question 3) Provide either a proof or a counterexample for the following statement. For integers a, b, and c, if a

divides bc, then either a divides b or a divides c.

Question 3 of interview 1 is particularly unusual in that three of the four students believed the false statement to be true, and the fourth (Amy) was unsure if the statement was true or false. None of the students recognized that this question was the definition of prime that Dr. S provided to the students, although Amy stated that she needed a composite number to find a counterexample. Amy is the only student who seemed to reflect on the problem and her examples in any manner. Amy chose values for three constructions (see Figure 19). The first construction was an example using the composite value 12 for the divisor. When she realized the construction was an example of the statement, she continued to construct a counterexample using the same divisor. However, Amy was not convinced, perhaps because the numbers she chose for the factors $b = 2$ and $c = 6$ were both less than the divisor $a = 12$. She continued to construct an additional example using $a = 3$ as the divisor. Although this was not useful as a counterexample, Amy thought to try an example with different properties, instead of trying the same divisor again. After considering these three constructions, Amy was not sure if the problems was true or false (even though she had a counterexample), so she decided to attempt a proof. Amy appeared to reflect on her progress, although her reflections were inconclusive. She lacked the experience and expertise to understand the implications of her work, although at the end of the interview during the reflection she did eventually conclude that the statement was false because of the counterexample she constructed.

Mike seemed to reflect some on his work. He was the only student to begin with a proof attempt instead of with examples. He started with proof by cases framework, but did not continue beyond the first case. After this proof attempt failed, Mike constructed a single example $2|(2 \cdot 3)$. This led Mike to believe the statement was true, and attempt another proof. Unfortunately, in this attempt he proved the converse, not the actual statement. Mike did not seem to reflect at all on the values he chose for his example, in particular the effect of a prime divisor. He may have reached this conclusion more quickly because he already believed the statement to be true before he constructed the example.

Raul and Carl included very little reflection on their work. Raul and Carl created two constructions each to try and decide if the statement was true or false. They both chose to use the same divisor on the examples they constructed. Raul considered $2 \nmid (3 \cdot 5)$ and $2|(4 \cdot 3)$, and Carl considered $4|(16 \cdot 4)$ and $4|(12 \cdot \quad)$. Carl did not finish his second construction because he realized the construction satisfied the statement no matter what value was selected for c. Neither of these students seemed to reflect on their initial constructions beyond a simple evaluation of whether the construction was a counterexample. Raul did not seem to consider the hypothesis or conclusion when selecting the values for his constructions. In fact, in all of his constructions during this interview, he listed values and then evaluated those values, rather than selecting values based on the hypothesis. Raul did not seem to reflect on his choice of a prime divisor at all, even after his proof by contraposition attempt failed. Similarly, Carl failed to reflect on his constructions, and in particular he did not recognize that he should not choose a b that is a multiple of a if you are seeking a counterexample. Carl failed to finish the second example because he recognized that the example would be true for any value of the third parameter, but he did not realize that was a consequence of his choices. Additionally, in the previous problem Carl had attempted six different counterexample constructions and was frustrated that none of those attempts produced a counterexample. His frustration seemed to carry through to his work on this problem, so he may have given up his search for a counterexample more quickly because of the emotion. Carl believed that his proof by contraposition was successful, but he applied the multiplication law of equality to an inequality using $\neq$, which is a logical error.

> **Task** (Interview 3, Question 4a) Prove or give a counterexample: All fine functions are periodic.

The students seemed to understand that constructing examples was insufficient proof for most statements. In all of the student interviews, there was only one instance of a student using examples as proof. During question 4a of interview 3, Amy used examples to "prove" that all fine functions are periodic. It appears that Amy reached this conclusion due to an insufficient understanding of the terms fine and periodic. In particular, the only fine functions that Amy could think of were transformations of $\sin x$ and $\tan x$, and she concluded that fine functions were trigonometric. Additionally, after reading the definition of periodic, Amy felt that the definition of periodic did not match her intuitive understanding, but accepted that the definition must be correct. In her attempt to prove that all fine functions are periodic, Amy argued that fine functions are periodic, because fine functions are trigonometric and trigonometric functions are periodic. She continued the argument by saying $\sin(\pi) = \sin(3\pi)$, and the period of sine is 2π, see Figure 22. From this evidence, Amy incorrectly concluded that all fine functions are periodic. Later in the interview, Amy used this statement in another proof. Since she had spent more time working with fine functions at this point, I asked her to question the accuracy of the assumption that all fine functions are periodic. At that point she constructed a counterexample to the statement, but she never returned to her initial argument.

One possible factor influencing Amy's behavior on this task is that she was visibly tired during this interview. This interview was conducted during the last

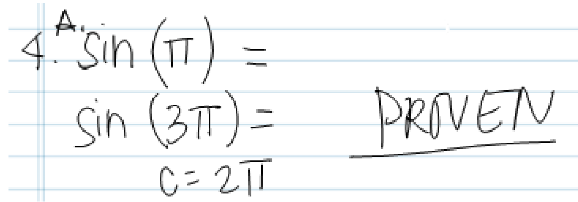

Figure 22: Amy's proof of question 4a of interview 3.

week of classes, and Amy seemed overexerted. It appears that her confusion about the definitions involved and her physical condition led her to overgeneralize. As a group, the students recognized that a single example does not prove a universally quantified statement and no such arguments were offered, except this single case.

Connecting to the language of formal proof. The students seldom connected their examples to their proofs. Typically the students used examples to help understand the definitions or to decide if the statement was true. When they decided the statement was true, they would try to write a proof. However, the students usually did not refer back to the example while writing the proof. For a student's example to be coded as connecting to the language of formal proof, the student had to refer to their example while writing formal language that could be a part of a proof. One reason this behavior received few codes was because students could complete problems without utilizing the technique.

> **Task** (Interview 3, Question 4b) Prove or give a counterexample: The product of a fine function and any other function is a fine function.

Carl was the only student who wrote a complete and correct proof and referenced an example in the process. He did so while working on question 4b of interview 3. Carl constructed the example $x^2 \sin x$ and determined from this example the key feature that the zeros of $\sin x$ make the product zero at those points. For the formal proof Carl wrote the general equation $f(x) \cdot g(x) = h(x)$, but talked about $f(x)$ being like $\sin x$ and $g(x)$ being like x^2. This was the only time Carl referred to an example during the actual writing of a proof, but it was also the best instance of this behavior from the students.

Task (Interview 1, Question 1) Let a, b, and c be natural numbers and $\gcd(a, b) = d$. Prove that a divides b if and only if $d = a$.

Raul also referenced an example while writing a proof, but he was only successful on one direction of a bidirectional proof. Raul constructed the example $\gcd(2, 4) = 2$ while working on question 1 of interview 1. From this example, Raul realized that the statement $d = a$ means that a has all of the properties of the greatest common divisor. From this, Raul was able to prove that $d = a$ implies that $a|b$. Raul and Amy were the only two students to realize that $d = a$ and the uniqueness of the greatest common divisor means that a is also the greatest common divisor. Neither Carl nor Mike made that observation. Additionally, neither Raul nor Amy realized that the opposite argument was also true, if a has the properties of the greatest common divisor, then $a = d$. In fact, all four of the students tried to use the properties of $d = \gcd(a, b)$ to reach the equation $d = a$. None of them tried to show that a had the properties of the $\gcd(a, b)$. Surprisingly the students who made the correct observation for one direction of a bidirectional claim, did not realize an analogous argument held for the converse of the statement.

Task (Interview 3, Question 1) Prove or give a counterexample:

a. If f and g are decreasing functions on an interval I, and $f \circ g$ is defined on I, then $f \circ g$ is decreasing on I.

b. If f and g are decreasing functions on an interval I, and $f \circ g$ is defined on I, then $f \circ g$ is increasing on I.

Both Amy and Raul referenced examples while working on question 1 of interview 3. Both students struggled with the definitions of increasing and decreasing functions. Amy had to look up the definitions and drew a sketch illustrating each term to make sure she understood the formal definition. Then Amy translated the statement of part b into symbols when she wrote $x < y$ and $f(g(x)) < f(g(y))$. Ultimately, Amy found counterexamples to both statement, however, she used an example to understand the definitions, which provided a framework for a possible proof. Raul also used examples to understand the definitions, however his approach was reversed by taking examples that he knew to be increasing or decreasing specifically $f(x) = 2x$ and $f(x) = -2x$ and determining if the definitions applied. He also used the understand he gleaned from the examples to write the framework for a proof, but did not complete the proof.

> **Task** (Interview 3, Question 4c) Prove or give a counterexample: Let $f : \mathbb{R} \to \mathbb{R}$ be a fine function. Let $g : \mathbb{R} \to \mathbb{R}$ be a function defined by $g(x) = f(x - k)$ for some $k \in \mathbb{Z}$. Prove that g is a fine function.

The final instance of using an example to connect to formal proof language occurred in Amy's third interview working on question 4c. As discussed in previous sections, Amy struggled greatly with the definition of a fine function. Her limited example space led her to erroneously claim that all fine functions are periodic; a fact which she used in her initial proof of this task. Additionally, Amy was exhibited signs of frustration at this point in the interview. As such, I provided additional hints on this question to allow her some success and to permit data from the remainder of the interview. I mention this now, so this evidence is weighed appropriately.

After Amy's initial attempt at this question, where she mistakenly used the assumption that all fine functions are periodic, I prompted her to consider other fine

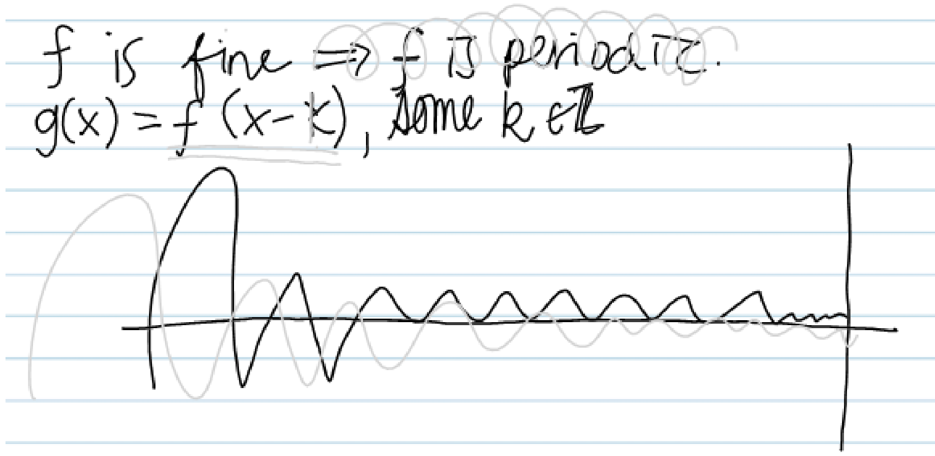

Figure 23: Amy's sketch to understand question 4c of interview 3. Her initial graph is in black, and the horizontally translated graph is in gray. This image also shows Amy's reconsideration of her assumption that all fine functions are periodic, as she wrote the statement (seen in black), but later crossed it out (in gray).

functions, and to question whether the assumption that all fine functions are truly periodic. At this point, Amy realized her previous assumption was false. She drew a sketch of a diminishing oscillating function, see the black graph in Figure 23, and observed that $g(x) = f(x - k)$ means f is being horizontally translated. Amy followed this with a sketch of the same graph horizontally translated, the gray graph in Figure 23. However, Amy continued to have trouble finishing this proof, specifically showing that the zeros would still occur at the integers. At this point, I provided one more intervention by asking Amy to consider the case $k = 1/2$. She argued with me that this was not possible because k must be an integer, but when I insisted Amy realized that the condition that k is an integer is required because without that condition the zeros will be moved off of the integers.

In this instance, Amy likely would not have considered the examples of the diminishing oscillating graph or the parameter $k = 1/2$ without my intervention. However, the connections be made between the graph and the proof was entirely her own behavior. She is the one who observed that the equation was a horizontal transformation of a graph, and she was also the one to realize the role of the closure of the integers. In both of these instances, Amy connected the language needed to the proof to the examples chosen successfully.

In general, these students seldom referred to their examples while writing their proofs, which is not surprising. Yet, when they made those references they were usually successful (at least partially, if not completely). Overall these particular students seem to able to use example successfully, however, they often are unable to reach any conclusions after constructing examples.

Dr. S's modeling. Dr. S only modeled implications of examples in a few circumstances. She tended to discuss the implications more frequently than she modeled it, and these discussions are elaborated in the next section. Primarily, Dr. S wanted the students to reflect on their work every step along the way. For instance, if a student started with a proof and got stuck, she wanted them to use an example to probe at the point where they got stuck in the proof. Then she hoped they would think about the properties in that example to be able to continue with the proof. Dr. S articulated this in an interview, when she said "I want them moving from example to proof. I don't want them stuck in the example mode, where they can't write a proof, but I don't want them stuck in the theoretical mode, where they're writing down words they don't know I'd like them to be bouncing back and forth." Overall, the goal is that the students will reflect after each attempt and use their accumulated knowledge to decide how to proceed next.

This reflection process has two components, starting with a proof attempt and using a construction to find the flaw or missing component, or starting with a construction and making deductions from that object to support the proof writing. The deductions possible from a construction include: that an object with the desired characteristics exists, that a conjecture *may* be true (in the case of an example), and that a conjecture is false (in the case of a counterexample). Additionally, an example may provide insight into the properties that cause the statement to be true.

With regards to the first component of reflection, Dr. S encouraged the students to think about how and why their proofs broke down and to use examples to reveal the errors in their proofs. Throughout the semester, Dr. S would reveal logical inconsistencies in the students' work by providing counterexamples. Often these inconsistencies occurred due to not thinking carefully about properties of the real numbers, e.g. one student claimed $n^2 > m^2$ means $n > m$, and Dr. S provided the class the counterexample $(-3)^2 > (-2)^2$, but $-3 \not> -2$. Dr. S also encouraged the students to reflect on their own proof attempts, and try to discover their own inconsistencies.

When starting with a construction, Dr. S modeled how to pay attention to details in proofs, especially with regards to quantifiers. During the first weeks of the semester, Dr. S spent approximately two hours of class time discussing quantifiers, negations of quantifiers, and how examples inform different types of quantifiers. For instance, on the second class day, Dr. S introduced several types of quantifiers. A simple statement instantiated each type of quantifier, see Figure 24. Each statement was evaluated for truthfulness, which led to either a proof or a counterexample to each statement, as seen on the right hand side. These examples introduced the role of reflection when evaluating truthfulness.

Another opportunity to talk about the importance of quantifiers occurred when Mike presented his solution to the task below.

> **Task** Consider the relation $\mathcal{S}$ on $\mathbb{N}$ defined by $x\mathcal{S}y$ if and only if $3|(x+y)$. Prove $\mathcal{S}$ is not an equivalence relation.

Before reading Mike's work, Dr. S discussed the negation of the equivalence relation

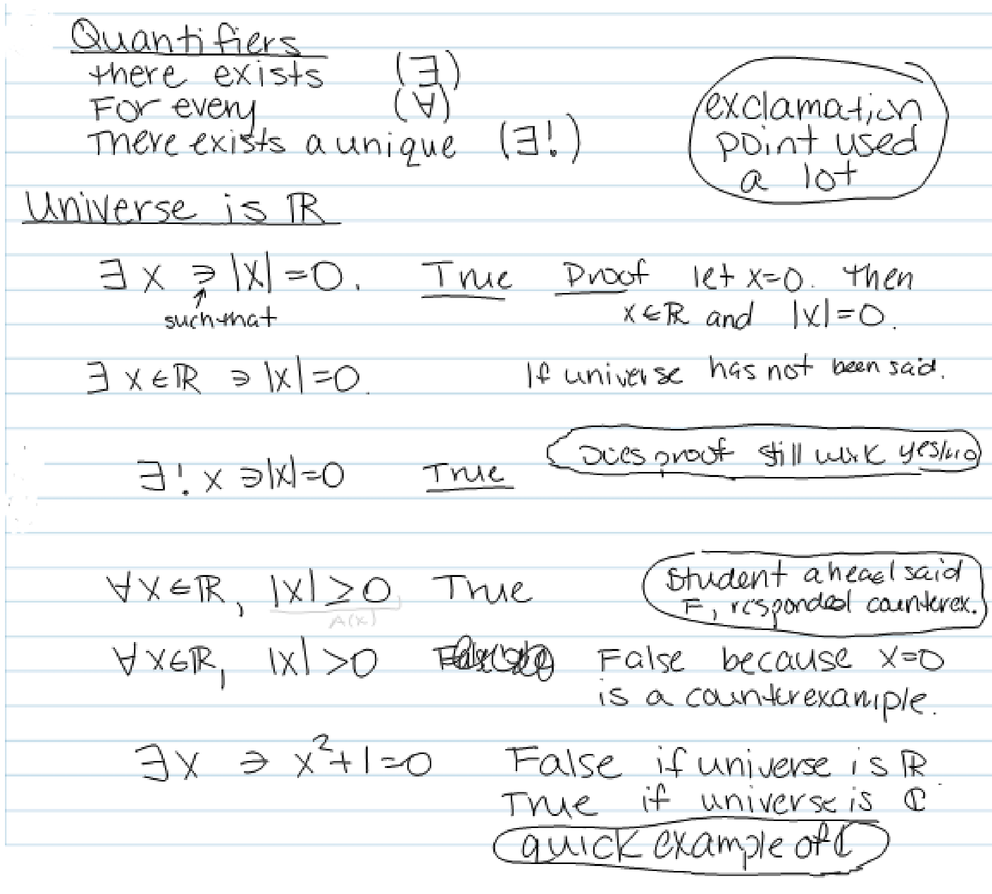

Figure 24: Dr. S provided the following statements to instantiate the various types of quantifiers. After describing the meaning and notation for each quantifier, Dr. S led the class through evaluating these statements. The image is from field notes from the day.

definition. Specifically, she used DeMorgan's Law to turn the "and" in the definition to an "or" in the negation. Thus she argued that a person only has to show that one of the three properties fails to show that the relation is not an equivalence relation. Continuing, the following dialog occurred

> Dr. S: How do you show a property fails?
>
> Carl: Uh ... Contradiction
>
> Dr. S: Contradiction is an excellent way sometimes
>
> Mike: Counterexample

> Dr. S: Counterexample is my favorite way. Thats also the best one. So, you can often just find a real counterexample. Concrete. Numbers. This is the time that you actually get to grab your examples and numbers.

Dr. S then turned to Mike's work, where he had attempted a proof by contradiction. Although Mike argued that this relation does not satisfy the reflexive property, his argument used general variables and did not include the necessary quantifiers. Dr. S then wrote the counterexample 3 does not divide $2 + 2 = 4$, so S is not reflexive.

Dr. S continued the discussion of this problem and asked the students "Does anyone know if it is transitive? And how would we figure it out? What is your thought process for how to figure it out?" She suggested they try a couple examples to get a sense as to whether it is true, and constructed the examples below.

$$1 \mathcal{S} 2$$

$$2 \mathcal{S} 4$$

$$1 \not\mathcal{S} 5$$

Dr. S concluded by saying "So by testing an example to see what was going on, we are done." Through this episodes, Dr. S reminded students about the importance of quantifiers and the relationship between quantifiers and negations. Also, Dr. S reminded the students of the role of counterexample in disproving a universally quantified statements, and modeled that approach.

Additionally, there were a few times when Dr. S lectured on tasks where the statement may or may not be true. In most of these instances the statement was false, and Dr. S produced a counterexample of the statement. In the others instances Dr. S produced one or more examples of the statement, and then concluded that the statement is probably true, and followed with the proof. Dr. S showed how these examples and counterexamples can help a prover evaluate the truthfulness of a statement.

Another implication is connecting examples to formal proof language. In this study, this behavior is operationalized by explicitly referring to an example while writing a proof. Dr. S modeled this behavior very limitedly. One instance occurred when Dr. S defined an ordering relation (also known as a poset). Dr. S provided the class the relation on the real numbers of x is related to y if and only if $x \leq y$. After defining a few additional terms, including upper bound and least upper bound, Dr. S proved the task below.

Task If an ordering relation has a least upper bound, then it is unique.

While discussing these definitions and proving this theorem, Dr. S switched between the notation for a generic ordering relation and the relation less than or equal to creates on subsets of the real numbers. Dr. S argued that anytime you work with an ordering relation you can really be thinking about less than or equal to on the real numbers, and treated this example like a generic example.

In another instance, Dr. S was reviewing a turn-in homework problem that asked the students to prove that the sets $A_a = \{(x, y) \in \mathbb{R} \times \mathbb{R} | y = a - x^2\}$ is a partition of $\mathbb{R} \times \mathbb{R}$. The final subproof of proving this statement required showing that $\mathbb{R} \times \mathbb{R} \subseteq \bigcup_{a \in \mathbb{R}} A_a$. Dr. S wrote "Let $(p, q) \in \mathbb{R} \times \mathbb{R}$. Then" on the board, and turned to the class and asked the class "what do I know? Where do I want to go?" Carl answered that they want $(p, q) \in A_a$ for at least one a. Dr. S replied "well, if I got to be in one of them, I need to figure out which one. How do I do that?" After about fifteen seconds of silence, Dr. S wrote the ordered pair $(5, 3) \in A_?$ on the board, and asked "what parabola does this live on?" Amy determined that $3 = a - 25$, so $a = 28$ and $(5, 3) \in A_{28}$. From this the class (at least three students talking simultaneously) announced that $q = a - p^2$ and $a = p^2 + q$. From this Dr. S completed the proof showing that $(p, q) \in A_a$ when $a = p^2 + q$. Through this excerpt of the course, Dr. S guided the students through the process of seeking the general through a specific example.

In summary, Dr. S taught the students to use examples to evaluate whether statements were true or false and to pay careful attention to the connection between examples and quantifiers. Additionally, Dr. S wanted the students to reflect on the results of example constructions and failed proofs. Although this was modeled in limited amounts, the class included some instances of connecting examples directly with the language needed for a proof. However, she only modeled single iterations of the reflection process, although she did talk to the students about completing multiple attempts at each problem.

Dr. S's discussion. Throughout the semester, Dr. S talked about the implications of using examples while working on proof-related tasks. She frequently encouraged the students to reflect on what they learned from examples and from failed proof attempts. Other implications from using examples on proof-related tasks include deciding the truthfulness of statements, finding counterexamples, generalizing or conjecturing, and using the example to write a formal proof. Throughout the semester, Dr. S talked about all of these implications at least once.

Dr. S repeatedly talked to the students about not giving up when working on their homework assignments. On the first day of class she encouraged the students to get problems wrong, because the learning occurs through fixing the solutions. She also suggested that the students present the board work problems that they had problems completing, because that would give them the individualized feedback they need. Dr. S repeated this advice anytime she reviewed a presented solution that contained significant gaps or errors, and would praise the student for choosing to present that problem.

In addition to persevering by not giving up, Dr. S also talked about reflecting on the completed work at each stage of the process. However, Dr. S only modeled single iterations of the reflection process. To encourage the students, Dr. S shared an anecdote of her days as a student, when she had a professor who insisted on 17 pages of attempts on a proof before he would help a student in office hours, because after that many attempts the student would have gotten close to the correct proof at least once. Dr. S shared this story to give the students comfort that multiple attempts are normal, and an expected part of doing mathematics.

When introducing proof techniques, Dr. S returned to the topic of quantifiers and examples. Dr. S discussed that in existence proofs, a prover often constructs an example and the proof merely verifies that the conditions hold. She followed this with three constructive existence proofs, and emphasized that existence proofs are the only time an example is a sufficient proof. Additionally, she provided the students with sample non-constructive existence proofs. During this same class period, Dr. S also discussed universal quantifiers and reminded the students that examples are insufficient to prove such statements.

Essentially, Dr. S told the students that examples are sufficient for statements including *there exists* as the quantifier, but that they are not sufficient when *for all* is the quantifier. In other words, Dr. S tried to get the students to realize that for an existence proof constructing an examples means they are done, but for a universal proof it does not. During the same lesson, Dr. S also discussed how a single example is (usually) sufficient as a counterexample. One weakness of these particular lessons was the lack of connections between a single counterexample being sufficient and the use of the universal quantifier in the statement. Since the majority of the statements in this course were universally quantified, it was not a significant problem, however it is unclear if the students recognized this connection.

Dr. S also encouraged the students to attempt the construction of examples and counterexamples to assess whether conjectures are true or false. In particular, Dr. S wanted the students to reflect on the examples they construct, because "sometimes examples can provide ideas" for the proof. During the same discussion, Dr. S instructed the students to reflect on their proof writing by saying, "If you ever do a proof, and it's wrong, don't throw it out. Use what it tells you." With these statements, Dr. S helped promote a culture of learning from mistakes.

There were fewer opportunities to discuss implications from the use of examples later in the course, because so many of the tasks demonstrated during class and assigned on homework asked for proof that specific constructions are truly examples. For instance, one homework problem asked to proved that the relation R on $\mathbb{N}$ is an equivalence relation when mRn if and only if m and n have the same digit in the tens place. This question asks the student to prove that the given construction is an example, and as a consequence there are limited opportunities for the students to use examples.

Due to the time constraints in the course, Dr. S chose to focus on single iterations of the reflection process and then to discuss what multiple iterations would look like. She modeled each of the implications she wanted the students to make, but not how to string them together when needed.

Comparing the instruction and the students. During the first interview, Dr. S stated that she wanted the students to be doing their homework, determining correct proof techniques for the problems, setting up an outline for an appropriate proof technique, and most of all she wanted them to not give up. She expected that the students would "still make mistakes on some of the newer concepts to them", and she wanted them to get comfortable with making mistakes.

Dr. S argued that some of the students were "still so afraid of making a mistake that they won't try anything." She tried to tell these students "scrap paper is your friend. Try anything you want, you don't have to turn it in." These quotes indicate the culture of reflection that Dr. S tried to insinuate into the students in her class. She wanted the students to try something, and if it does not work to learn from it and try again.

At the time of the first interview, the students were hesitant to complete too much reflection. The students tended to reflect on what their examples told them about the truth of the statements, but not on the quality, effectiveness or construction of their examples. For instance, none of the students observed that question 3, which says "if $a|bc$, then $a|b$ or $a|c$", is the definition of prime given to them by Dr. S. Additionally, the students did not seem to reflect on their failed proof attempts, and how to improve them. In fact, when Amy attempted question 1 (about greatest common divisors), she essentially wrote the same incorrect proof twice, even though the two proof attempts occurred nearly an hour apart. Amy even used the same counterexample to disprove the sub-claim she formed.

The students did perform some basic reflection particularly using their constructions to evaluate if the statement is true or false. Since most of the examples were constructed when the students were seeking counterexamples, they usually decided to construct another example if their first construction was not a counterexample. Usually the students had constructed a few examples before deciding that a statement is true, and stopped once they had identified a counterexample. The students often drew no conclusion at all from their reflections.

The students remained fairly consistent with their level of reflection throughout the semester. They started the semester uncertain about the conclusions they can draw, and gained only a minimal amount of confidence in this aspect throughout the semester. The only implications that increased significantly over the semester was connecting an example to formal proof language, and drawing no conclusion at all. Additionally, the students typically only went through one or two iterations of reflections and reattempting.

Dr. S expressed a hope that the students would continue grow in the frequency in which they reflection on their proofs and examples. "I want them moving from example to proof, I don't want them stuck in the example mode where they can't write a proof, but I don't want them stuck in the theoretical mode where they're writing down words they don't know what mean. I'd like them to be bouncing back and forth." However, Dr. S only demonstrated one or two iterations of this process during the lectures. During the member checking interview, she said this was due to the time constraints in class, and the need to progress through the content. Dr. S wished she could have modeled more extended reflection processes, but that it was not possible. However, she confirmed that she would often talk to the students about more extended reflection during office hours when time is less constrained and fewer students are present at a time.

Overall, the students were limited in the conclusions they would draw after constructing examples. They usually concluded the truth of a statement or made no conclusion at all. In a few circumstance the students connected their examples to a formal proof. These implications are consistent with what was modeled in the classroom. Dr. S modeled reflecting on examples and proof attempts with one or two iterations, deciding truthfulness of statements, and a few circumstances of connecting examples to formal proof writing. Her instruction and the behavior of the students seem to overlap significantly.

Summary of the implications. The implications made by the students included: 1) constructing another example, 2) deciding the truthfulness of the statement, 3) making a conjecture, 4) connecting the example to a proof, and 5) doing nothing. All of these behaviors were demonstrated in the lecture, except doing nothing. Additionally, Dr. S discussed using examples to find logical flaws or gaps in arguments, and focusing on the role of quantifiers when considering examples.

Summary Tables

Overall, Dr. S taught aspects of all four components of effective example usage. Dr. S extensively modeled the indicators of when to consider examples and the purposes for constructing examples. In addition to modeling these behaviors, she also talked about when and why examples can be useful when working on proof-related tasks. Dr. S provided less instruction on example construction and the implications from using examples, especially connecting directly to the proofs. Dr. S expected the students to figure out some of these skills, particularly example construction for themselves.

Amy constructed examples repeatedly during her interviews. Amy was particularly good at monitoring her thinking, and constructing examples when she needed insight or to understand a statement. Nearly all of her examples were accurate constructions. However, Amy struggled with the implications from examples. In one instance she did not recognize when she had constructed a counterexample, and in another she considered a single example to be a proof. Amy was successful in recognizing indicators, knowing purposes and constructing examples, but not with the implications from using examples.

Carl exclusively used examples on tasks with language that indicated the appropriateness of examples, including the phrases *prove or disprove*, *make a conjecture*, *generalize* or *construct an example*. He also typically only used one example per task, and most of the examples had the purpose of determining if the

Table 8: A summary of Amy's example use during the three interviews.

Indicator	No.	Purpose	No.
Prove or Disprove	17	Understand a Statement	15
Make a Conjecture	4	Evaluate T/F	16
Seeking Insight	10	Make a Conjecture	3
Construct/Verify Example	2	Generate Proof	2
Generalize	3		

Implications	No.
Construct Another Example	15
Decided Truthfulness	8
Connect to Formal Proof	4
Made a Conjecture	3
Did Nothing	6

Construction Accuracy	No.	Construction Technique	No.
Accurate Construction	30	Authoritarian Source	3
Missing Hypotheses	1	Trial and Error	17
Missing Conclusion	0	Transformation	16
Incomplete	5	Analysis	0

statement is true or false. Most of his constructions were accurate, even though he primarily used a *trial and error* technique. Generally, Carl drew accurate conclusions after constructing examples. However, this is likely due to his reluctance to construct examples. Overall, Carl was successful with using examples, but his use of examples was limited to very specific instances. It is unclear whether Carl would have the same success in the wider variety of purposes.

Raul was a moderate user of examples. He used examples regularly to improve his understanding and based on the language of the tasks presented, but he did not use them as frequently as Amy. Raul was also successful with determining the implications of the examples. Although there were a few times that Raul did not reach any conclusions from the examples, he generally was able to use his examples to decide the truth of statements and to connect to the proof. Raul's

Table 9: A summary of Carl's example use during the three interviews.

Indicator	No.	Purpose	No.
Prove or Disprove	15	Understand a Statement	3
Make a Conjecture	1	Evaluate T/F	15
Seeking Insight	0	Make a Conjecture	1
Construct/Verify Example	2	Generate Proof	0
Generalize	1		

Implications	No.
Construct Another Example	6
Decided Truthfulness	6
Connect to Formal Proof	2
Made a Conjecture	1
Did Nothing	4

Construction Accuracy	No.	Construction Technique	No.
Accurate Construction	16	Authoritarian Source	2
Missing Hypotheses	0	Trial and Error	15
Missing Conclusion	1	Transformation	2
Incomplete	2	Analysis	0

biggest struggle with examples was with construction, although those problems only occurred during the first interview.

Mike used a limited number of examples because only two of the three interviews were conducted. Additionally, Mike spent most of the second interview on the proof validation task [Interview 2, Question 1], and as such only constructed one example during the entire second interview. Mike generally constructed examples that did satisfy the necessary conditions, however, his examples were generally poorly selected to achieve his desired goals. Mike seemed to understand the value of reflection when working on proof-related tasks, but was generally unsuccessful in conducting such reflections. One potential reason for this lack of success on the implications is that it was not possible to observe Mike's growth throughout the semester.

Table 10: A summary of Raul's example use during the three interviews.

Indicator	No.	Purpose	No.
Prove or Disprove	7	Understand a Statement	10
Make a Conjecture	3	Evaluate T/F	9
Seeking Insight	8	Make a Conjecture	1
Construct/Verify Example	3	Generate Proof	1
Generalize	0		

Implications	No.
Construct Another Example	6
Decided Truthfulness	6
Connect to Formal Proof	3
Made a Conjecture	1
Did Nothing	5

Construction Accuracy	No.	Construction Technique	No.
Accurate Construction	18	Authoritarian Source	2
Missing Hypotheses	3	Trial and Error	11
Missing Conclusion	0	Transformation	8
Incomplete	0	Analysis	0

The instructor predictions of student choices on interview tasks compared to the student responses. In the final interview, Dr. S reviewed the tasks given to students during their interviews and tried to predict whether the students would consider examples. The results are found in Table 12. There was only one instance when Dr. S thought that the students would use examples and they did not (2.5), and only one student even attempted this question. There were times that the students used examples when Dr. S did not think they would. However, in general her predictions were accurate.

Overall, Dr. S seemed to have a realistic expectations of the capabilities of students at this level. During the member checking interview, Dr. S mentioned that when she first started teaching this course her expectations were significantly different, but over her years of experience her expectations have adjusted. Much of

Table 11: A summary of Mike's example use during the three interviews.

Indicator	No.	Purpose	No.
Prove or Disprove	3	Understand a Statement	3
Make a Conjecture	0	Evaluate T/F	3
Seeking Insight	3	Make a Conjecture	0
Construct/Verify Example	0	Generate Proof	0
Generalize	0		

Implications	No.
Construct Another Example	2
Decided Truthfulness	3
Connect to Formal Proof	0
Made a Conjecture	0
Did Nothing	1

Construction Accuracy	No.	Construction Technique	No.
Accurate Construction	4	Authoritarian Source	0
Missing Hypotheses	1	Trial and Error	4
Missing Conclusion	0	Transformation	2
Incomplete	1	Analysis	0

the overlap between Dr. S's instruction and the student's behavior can probably be attributed to her realistic expectations of the students.

Table 12: This table shows the predictions made by Dr. S and how many students used examples compared to how many attempted the problem.

Int.Task	Dr. S's prediction	Ex:Att
1.1	Weaker students may use examples	3:4
1.2	Start with examples, unless known	3:4
1.3	Start with examples	4:4
1.4	Probably not	0:3
2.1	Maybe when proof fails	2:4
2.2	No	0:2
2.3	Maybe	2:2
2.4	Start with 1337 and then small n and m	2:3
2.5	Yes, see what happens	0:1
2.6	No	1:1
3.1	Yes	3:3
3.2	Yes, see what's going on	2:3
3.3	Yes, new definition	3:3
3.4a	Yes, with a picture	3:3
3.4b	Maybe not	2:3
3.4c	Maybe	1:2
3.5	Impressed if they don't panic	3:3

CHAPTER 5

DISCUSSION

A Theory of Effective Example Use

Through the grounding of the data from both the students' and the instructor's point of view regarding the use of examples while working on proof-related tasks, four crucial phases emerged: indicators (when), purpose (why), construction (how) and implications (what now). In the instances when a student's use of examples was unsuccessful on a task it was because something failed in one or more of the four phases. For instance, Raul stated that he was seeking a counterexample (purpose), because the task said *prove or disprove* (indicator). However, Raul mistakenly identified a non-example as a counterexample (construction), which led to him reaching the incorrect decision about the truthfulness of the statement (implications). In this situation, Raul understood the purpose and the indicator for using examples, but he failed in the construction. In other situations, the student failed in the implication phase, and in a few instances the student failed in the indicator phase.

However, the students were successful in their examples use when they could articulate ideas or implement behavior that spanned all four phases. For instance, when Carl proved that the product of a fine function and any other function is fine, he observed the statement said *prove or disprove* (indicator), he constructed an example $x^2 \cdot \sin x$ (correct example, if one assumes the correct transformation of $\sin x$), he recognized that this example helps him evaluate truth and reveals an underlying structure (purpose) and he used this example and its structure to write an accurate proof (implication). From these observations of student success and failures, the framework of the four phases for effective example usage emerged.

This seems to be the first time that the entire arc of using examples during proof-related tasks has been studied. Previous studies have typically focused on one type of proof task and one phase of the process (Alcock & Simpson, 2004, 2005; Alcock & Weber, 2005, 2010; Iannone et al., 2011; Moore, 1994).

As data analysis proceeded to look at the instruction, the initial framework still seemed to apply. When Dr. S commented on the use of examples, she tended to emphasize one or more of the four phases: indicators, purposes, constructions and implications.

Indicators for Using Examples

As seen in the results from both the instructor and the students, the indicators that examples might be useful for the problem solving process fell into three categories: learning a new definition, prompts from the language of the task, and recognizing a need for additional insight via self-monitoring.

The students had mixed results with recognizing the indicators of examples use. The generally were extremely consistent with the indicators from task language, which may be a result of previous mathematics instruction. The students may be used to a procedural problem-solving curriculum (from the algebra and calculus sequence), where *keywords* often indicate the correct technique to solve the problem. Lester and Garofalo (1982) observed this reliance on *keywords* as a common problem solving technique among elementary students, and it may be surmised that this strategy will persist. Furthermore, task language indicators frequently provide clues towards the intended purpose of the examples. A task with the directions *prove or disprove* indicates that examples may be useful for evaluating the truth of the statement. A task with the direction *conjecture* indicates that examples may be useful for determining a pattern and making a conjecture.

So, not only do these indicators suggest the use of examples, but they also provide insight into the intended purpose of these examples.

The students had more trouble with the self-monitoring indicators for using examples. The instructor was unable to genuinely demonstrate how to monitor ones thinking for confusion, as she already had an understanding of the definitions and statements presented in the course. However, it seemed as though some of the students did not have the maturity to recognize what they do and do not understand, and furthermore to recognize when example could provide the insight they were seeking.

Constructing examples of new definitions was the most prevalent use of examples from the instructor. There were many classes where the only examples presented were examples of new definitions. However, the students did not always recognize that they should construct examples of unfamiliar definitions. In the third interview, two new definitions were presented: fine function, and periodic function. Fine function was a completely new definition which the students would never have seen before, and one of the tasks explicitly asked the students to construct one or more examples of this definition. Since it was a stated task, all of the students who completed the third interview did construct at least one example of a fine function. Periodic function was likely a familiar definition from previous coursework, but had not been discussed in this course, so the rigorous definition was new to the students. Amy was the only student who recognized that this was a new definition, and who rigorously showed an example of the definition. It is unclear why the other students did not construct examples of this definition, but possible explanations include: they already understood the definition from previous classes, or they did not realize that the definition was new.

Indicators have not been studied in the literature on examples and proof writing. The studies on semantic and syntactic proof productions hypothesized that personal preference is the primary indicator of whether an individual will use semantic or syntactic reasoning primarily (Alcock & Inglis, 2008, 2009; Weber & Alcock, 2004). These results seem to partially contradict this assumption. Personal factors, such as the ability to self-monitor, impact whether a prover decides to use examples on a task. However, these are not the only indicators for using examples, and in this data set the majority of examples were inspired by task language, not by personal factors. Two explanations for the frequency of the task language indicating example use are that the students latched onto the key phrases in the task language, like the third and fifth graders in Lester and Garofalo (1982), or simply most of the tasks contained the observed phrases.

Purpose of Examples

The purposes for constructing examples consisted of six categories: understanding a statement, seeking connections between definitions, determining the truth of the assertion/generating a counterexample, generating a proof, making a conjecture, and evaluating arguments/revealing logical inconsistencies. The students could generally recite the purposes for using examples, especially understanding a statement and evaluating the truth of an assertion. In several instances, the students could recite the purpose of using examples, but did not actually achieve the stated purpose (see more in the implications section).

Previous research on the use of examples in proof writing, revealed four categories: understanding a statement, evaluating the truth of an assertion, generating a counterexample, and generating a proof (Alcock & Weber, 2010). In this study, counterexamples were used to establish that a statement is false, and to evaluate arguments. As such, it seemed appropriate to revise Alcock and Weber's *generating a counterexample* category to separate the two uses of counterexamples.

However, evaluating the truth of an assertion includes generating a counterexample to disprove a statement, so those categories were merged. Evaluating arguments was not a component of Alcock and Weber's study because they limited their tasks to just writing proofs, not the broader proof-related tasks used in this study.

Similarly, the category *making a conjecture* was another category that arose from the broader types of tasks. A common technique for making a conjecture is generating several examples in order to observe a pattern. However, if a prover is not being asked to make a conjecture, this purpose will not be revealed in the data set. This is why this purpose was not included in Alcock and Weber's (2010) framework for the purposes of examples in proof writing.

One observation regarding the purposes of examples is that the instructor and the students often used examples in different ways in the classroom environment. The instructor primarily used examples to improve the student's understandings of definitions, and to reveal logical inconsistencies in (the students') arguments. When presenting a new definition, or when the students demonstrate confusion about a statement, Dr. S would present examples that probe at the subtleties of the statement. For instance, Dr. S presented approximately a dozen different relations when she defined relation and equivalence relation. She did this to introduce the variety of notations that are possible for relations, and to provide some examples that satisfy all, some, or none of the conditions for being an equivalence relation. She recognized that these definitions were notationally complex and that the students would need several examples to learn the details of the definitions.

Using examples or counterexamples to evaluate arguments was another frequent use of examples by the instructor. This was a consequence of the students' presentations on the blackboard at the beginning of class, where the students often presented work containing errors. Dr. S tried to encourage the students to present solutions to the problems that they found the most difficult, because this was an opportunity for the students to get corrective feedback without it negatively affecting their grades. This provided Dr. S many opportunities to correct their arguments, and to reveal the errors that they made.

The students almost exclusively used examples to evaluate the truth of an assertion or to understand a statement. This is partly due to the frequency of the types of tasks presented to the students. For instance, the majority of the interview tasks included the phrase *prove or disprove* which led to the students using examples for the purpose of evaluating the truth of the assertion with the greatest frequency. The students seldom used examples with the intended purpose of generating a proof, however, there were instances where the students intended to use the example to evaluate the truth of the assertion, but ultimately use the example to evaluate truth and also generate a proof. This is a key difference between the purpose and the implication phases: the purpose addresses why a prover might choose to use examples and the implications is what a prover actually concludes from using examples.

When the students used examples to understand a statement, approximately half the instance occurred to improve understanding of a definition and the other half occurred to improve understanding of the problem statement. This contrasts with the instructor who used examples to improve understanding of definitions with far greater frequency than understanding proof statements. This is likely are artifact of the different environments, meaning the instructor was introducing many new definitions, whereas the students only experienced one brand new definition during the interview.

Construction of Examples

The construction of examples is the phase of this process in which the students showed increased sophistication throughout the semester. During the first interview, several of the students misidentified constructions as examples or counterexamples. Furthermore, two of the students used the weakest form of *trial and error*, where they randomly picked values for the integers in the problem statement and then tested the hypotheses and conclusions. The other two students used a more sophisticated form of *trial and error*, where they randomly picked one value, but then tried to pick meaningful values for the remaining variables. By the end of the semester, all of the students were selecting examples with more thought, even if they were still using a *trial and error* technique.

Additionally, the students used the *transformation* construction technique with increased frequency throughout the three interviews. During the first interview, Mike was the only student to use the *transformation* technique. However, by the third interview, the students used the *transformation* technique for more than half of the examples. It is unclear exactly what caused this growth. Possible explanations include the students' individual development throughout the semester, the influence from the instruction, and the new content.

It is logical to expect the students to learn new techniques for example generation throughout the semester. Although example generation is not the primary objective from this course, it is a secondary objective. The board work homework problems frequently included example construction problems, and Dr. S spent class time constructing and presenting examples especially later in the semester. Dr. S expressed the hope that the students would move to more advanced techniques of example construction as they progressed, but argued that a student had to be "ready for that." It appears that the students in the sample achieved this level of readiness throughout the semester.

The mathematical content of the tasks (and the examples) changed throughout the semester and matched the recent content from the course. The first interview focused on number theory topics such as divisibility and greatest common divisors, the second interview focused on sets and relations, and the third interview focused on functions on the real numbers. One possible explanation for the increased use of the *transformation* example construction technique is that transformations of real valued functions is a familiar topic from the algebra-trigonometry-calculus sequence. As an illustration, all of the students who completed the final interview talked about how $\sin(x)$ can be transformed into an example of a fine function. It appears that the content may have played a role in the students' choice of construction technique.

Previous research on undergraduate example construction showed that the students used *trial and error* techniques approximately 80% of the time (Iannone et al., 2011). This percentage is considerably higher than than the 57% *trial and error* observed in this study. It is unclear what accounts for this discrepancy, although the most likely causes are the sample or the task selection. The students in both studies were in their second or third year of their studies, although the students in Iannone et al. (2011) were from the UK, and likely had more mathematics preparation than those in this study. Both studies also had small samples, this one had four participants and Iannone et al. (2011) had nine. As such, the individual characteristics of the participants could strongly affect the percentages. Additionally, The two studies did use different tasks. Iannone et al. (2011) provided the students with the definition for a real-valued function f that is preserved on a subset A, meaning $f(A) \subseteq A$, and a series of construction tasks involving this definition. This was an unfamiliar definition for the students, which may have contributed to their decision to use *trial and error*. However, in this study, when the students were presented with the definition for a fine function, which was also an unfamiliar definition concerning real-valued functions they tended to use *transformation* instead of *trial and error*.

This study and Iannone et al. (2011) were similar in that none of the students used an *analysis* construction technique. This result is consistent with Antonini (2006) who believed that the *analysis* technique is only used by expert mathematicians when the other techniques have failed. Since the students in this study are novices it is unsurprising that they never attempted an *analysis* technique for constructing examples.

Implications of Examples

The students did better with the more procedural implications (construct another example, decide truthfulness), but struggled with the more open ended implications (connect to proof language, and make a conjecture).

The students frequently constructed examples to evaluate the truth of tasks with the directions *prove or disprove*. Frequently the students reached the decision that a statement was true after constructing only one or two examples. Although the students were generally correct with their evaluations, in a few instances, particularly in question 3 of interview 1 this led to the phenomenon of overconfidence as described by Buchbinder and Zaslavsky (2011). On this particular task, three of the students believed the statement to be true from the construction of at most two examples. After concluding that the statement was true, two of the students went on to write arguments that they believed to be proofs, but that used facts about divisibility that are not true.

Alcock and Weber (2010) showed that undergraduate students have trouble connecting examples to formal proofs, similar to the struggles observed in the students in this study. In particular, Alcock and Weber (2010) observed that sometime undergraduate provers can used example to obtain the key idea for a proof, but feel lost when trying to turn this statement into a formal proof. This is precisely what happened to Carl during the first interview, when he constructed several examples, observed a pattern, but did not know how to turn that observation into a proof. (See the discussion around Figure 21 for more details.)

In general, the students made appropriate conclusions from the constructions they created. The students knew that an example does not prove a universally quantified statement, and there was only one instance where an example was presented as proof of the statement. The students used examples to evaluate whether conjectured statements are true or false, and to gain insight into the definitions, statements, and properties needed for the argument.

The students did not always persevere when working on the interview tasks. In several instances, a student would give up on a task before they were satisfied with their solution. A few times this was due to running out of time for the interview, but more frequently they did not know how to proceed. Typically at this point they would simply give up, although in a few instances they would try another approach.

Modeling Behavior versus Discussing Behavior

Dr. S modeled some behaviors, but only discussed others. For instance, Dr. S only modeled single iterations of bouncing between examples and proof attempts, but she talked about completing multiple iterations when needed. She also frequently modeled the indicators and purposes for using examples, but tended to only discuss construction techniques and implications.

The students appear to have better success with the phases that were modeled with the greatest frequency, namely the indicators and the purposes of examples. All of the students could recite purposes of examples, especially the role of examples in evaluating truthfulness, which was frequently discussed and modeled by Dr. S. However, the phases that were the most challenging for students (construction and implications) were the least frequently modeled by Dr. S.

There is insufficient data to claim that the frequent modeling of the indicator and purpose phases caused the behavior of the students; however, it must be a contributing factor. It is also possible that the behaviors in these phases are easier to implement, so resulted in more frequent use from both the instructor and the students. The construction and implication phases involve more opportunity for creativity and personal choice. As such, these phases could be more difficult to teach and learn, which might explain the low frequencies in the lecture and the difficulties experienced by the students.

Implications for Teaching Transition-to-Proof Courses

One of the important implications from this study is the importance of teaching students to address details when working on proof-related tasks. Students need to be explicitly taught the relationships between universal quantifiers, existential quantifiers, examples and counterexamples, consistent with the recommendations of Epp (2003). If students recognize that counterexamples are the negation of a universally quantified statement, this will help them remember the conditions that a construction must satisfy to be classified as a counterexample. Although this class included a lecture on negating universally quantified statements, the connection between such negations and counterexamples was not explicitly stated. This may partially explain some of the difficulties the students had with evaluating whether constructions are actually counterexamples during the first interview.

Another implication is that students should be explicitly taught strategies for constructing and verifying examples. One of the hardest parts of trial and error is picking the construction to test. However, by explaining how the examples in the course are constructed, it may be possible to guide the students beyond blinding picking parameters to test.

In this study, most of the students became convinced that a *prove or disprove* statement was true after constructing only one or two examples. However, when mathematicians obtain conviction from empirical evidence it is often from multiple examples or for unusual properties (Weber, 2013; Weber et al., 2014). Although it is unreasonable to assume that numerous examples should be constructed before trying to prove a statement, we need to teach students to consider the quality of the examples they construct, and to view the examples as a collection. For example, a statement that is true for a prime number, a perfect square, and another composite number is far more believable than a statement evaluated only with a prime number. But students need to be taught to consider examples collectively rather than individually.

Instructors should model how to use examples generically, and how to consider the properties within examples. When first learning these skills (and later), it can be easy to become overwhelmed by the definitions and the construction of examples. But in order to use examples to generate a proof, a prover needs to be able to see the general properties within the specific example. Additionally, the prover needs to have sufficient familiarity with the language and notations needed for proof. Students who see this type of reasoning modeled repeatedly will be able to implement it more effectively than those who do not.

Examples of Proof Types

For this study, sample proofs were purposefully excluded from the definition of an example. The primary purpose of this course is to teach students how to write proofs, especially how to set up different types of proofs. As a result, the instructor presented many different proofs that were instances of different types of proof, including direct proofs, proofs by contraposition, and proofs by contradiction. A particular direct proof could be considered an example of a direct proof. These sample proofs were excluded from this study because using them to assist in writing other proofs would be a form of syntactic reasoning rather than semantic. These sample proofs can provide insight into how to set proofs up, but provide only limited insight into the definitions and concepts involved in the statement to be proved.

Although these examples were not the focus of this study, they could form the basis of another study. Students likely use the sample proofs in their notes to assist them on their assignments. However, this behavior has not been studied. Such a study would provide additional insight into how students use their notes and the role of sample proofs in transition-to-proof courses.

Future Research

Additional studies need to be completed to further verify the reliability of this four phase model. The model was grounded in the data from both the students and the instruction, but it has not been tested with other populations or samples. Does this model continue to represent how to effectively use examples when we consider graduate students or mathematicians? Does it vary between pure and applied mathematics topics?

Other questions about this model concern the instruction on this topic. How does instruction impact a provers ability to effectively use examples? Does explicit instruction in these phases impact the ability to use examples effectively? Although there were aspects of all four phases in the instruction in this class, the phases were not taught explicitly. It is unclear whether or not such instruction will actually help the students learn how to use the examples effectively. Some studies suggest that instruction in problem solving frameworks alone does not help students become better problem solvers (Garofalo & Lester, 1985; Schoenfeld, 1980), so it is possible a similar phenomenon will occur here. This can only be established through additional testing and study.

Finally, it is unclear whether effective example use will positively impact proof writing. Iannone et al. (2011) found that generating examples provided no benefits to the students as compared to receiving a list of examples. One interpretation of this is that it does not matter where the examples come from, what matters is how the examples are used and what conclusions are drawn from the examples. As such, it is possible that knowledge in using examples effectively can improve a persons ability to successfully write proofs, but additional study is needed on this topic.

APPENDIX SECTION

APPENDIX A

IRB EXEMPTION

Based on the information in IRB Exemption Request EXP2013U286735G which you submitted on 11/19/13 15:35:22, your project is exempt from full or expedited review by the Texas State Institutional Review Board.

Based on the information in IRB Continuation Request CON2014J926, your request is approved.

APPENDIX B

CONSENT FORM

I have the opportunity to participate in the research study titled "Changes in Mathematical Thinking Regarding Proof" conducted by Ms. Sarah Hanusch from the Department of Mathematics at Texas State University. I understand that my participation is voluntary. I can stop taking part without giving any reason, and without penalty.

PURPOSE

The purpose of the project is to analyze the effectiveness of teaching proof techniques to undergraduate students enrolled in Introduction to Advanced Mathematics (MATH 3330). The primary goals of this project is to 1) observe how the instructor presented proof techniques, particularly the use of examples during proof, and 2) how the students utilize proof techniques in their own work. This project is the researcher's dissertation research.

PROCEDURES

Data will be collected through classroom observations, collection of classroom artifacts and interviews. The field notes from the observations and interviews may be supplemented by still photos, audio or video recordings.

BENEFITS

The benefit to me, as a participant, is an opportunity to express opinions regarding the course and to develop ideas more deeply. Some compensation may be available to interview participants. Participating in the study will not affect my grade in the course.

CONFIDENTIALITY

Any data gathered, including observation notes, audiotapes, copies of students' written work, and videotapes will be securely stored. The audio and video recordings will be used for research purposes, and may be included in research presentations. All recordings will be destroyed five years after the dissertation is completed. No information that identifies a participant will be shared with those outside of the researcher's dissertation committee.

FURTHER QUESTIONS

The researcher will answer any further questions about this research, now or during the course of the project. The primary contact person is Sarah Hanusch (sh1609@txstate.edu).

CONSENT

I consent to participate in this study.

Participant Name ________________________________(please print)

Participant Signature ______________________________ Date ______________

PHOTO/AUDIO/VIDEO RELEASE

I, ________________________________, agree to be photographed and videotaped for this research by Sarah Hanusch without receiving compensation of any kind. I understand that this footage may be used, as deemed appropriate by Sarah Hanusch, for future presentations which may be viewed by public and private sector audiences. If I do not consent, I will be edited out of all recordings.

Participant Signature ______________________________ Date ______________

APPENDIX C

INITIAL BACKGROUND SURVEY

Please provide some background information:

Phone ____________ **School email** ____________ @txstate.edu

Major ____________________ **Minor** ____________________

No. of Credits Completed ________________

Are you seeking a secondary teaching certificate? yes no

Approximate GPA 0.00-1.99 2.01-2.49 2.50-2.99 3.00-3.49 3.50-4.00

This is my ____________ (first, second, etc.) time taking MATH 3330.

Previous Mathematics Courses mark with the number of time you attempted each course and your grade in your last attempt. Also mark courses that you are currently enrolled in with an IP for in-progress.

Course	# of Times	Grade
Math 2472 (Cal II)		
Math 2358 (Discrete I)		
Math 3373 (Cal III)		
Math 3305 (Prob/Stat)		
Math 3377 (Linear Alg.)		
Math 3323 (Diff Eq)		
Math 3315 (Geometry)		
Math 3398 (Discrete II)		
Math 3325 (Number Systems)		
Math 3348 (Operations Research)		
Other		

Thank you for participating.

APPENDIX D

FIRST INTERVIEW PROTOCOL WITH THE STUDENTS

Initial Questions

These questions were asked before giving the students' the tasks.

1. What have you learned so far this semester?
2. Describe the goals of the course from the instructor's point of view.
3. What are your goals for this course? Do you think they are different from the instructors?
4. Do you think you are meeting the instructor's goals?
5. How do you define proof? What constitutes a proof?
6. What convinces you that a mathematical statement is true?
7. After you read a statement that requires proof, what is the first thing you do?
8. What are some of the other things you think about when starting a proof?
9. What are some of the strategies that you have learned this semester?

Tasks

The students were given approximately 45 to 60 minutes to work on the following problems. They were asked to show all scratch work, and to think out loud.

1. Let a, b, and c be natural numbers and $\gcd(a, b) = d$. Prove that a divides b if and only if $d = a$.

2. Provide either a proof or a counterexample for the following statement. For integers a, b, c, and d, if a divides $b-c$ and a divides $c-d$, then a divides $b-d$.

3. Provide either a proof or a counterexample for the following statement. For integers a, b, and c, if a divides bc, then either a divides b or a divides c.

4. Let a, b, c, and d be positive integers. Prove that ac divides bc if and only if a divides b.

Reflection Questions

After completing the task, students were asked to reflect on their work with the following questions.

1. Will you describe any strategies that you used while working on this task?

2. Why did you choose this strategy?

3. Do you think you implemented the strategy effectively? Explain.

4. What other strategies could you have used?

APPENDIX E

SECOND INTERVIEW PROTOCOL WITH THE STUDENTS

Initial Questions

These questions were asked before giving the students' the tasks.

1. What strategies have you learned so far this semester?

2. Do you think you are using all of the strategies that the instructor would like you to use?

3. I would like you to pull out your notes from lecture. Do your notes include the strategies and "helpful hints" that Dr. S talks about or do you just write what she writes? Can you show me some instances of strategies that you have included?

Tasks

The students were given approximately 45 to 60 minutes to work on the following problems. They were asked to show all scratch work, and to think out loud.

1. Assign a grade A (correct), C (partially correct), of F (failure) to each. Justify your assignment of grades.

   (a) Claim: If $X = \{x \in \mathbb{N} : x^2 < 14\}$ and $Y = \{1, 2, 3\}$, then $X = Y$.
   "Proof:" Since $1^2 = 1 < 14$, $2^2 = 4 < 14$, and $3^2 = 9 < 14$, $X = Y$.

   (b) Claim: If A, B, and C are sets, and $A \subseteq B$ and $B \subseteq C$, then $A \subseteq C$.
   "Proof:" If $x \in C$, then, since $B \subseteq C$, $x \in B$. Since $A \subseteq B$ and $x \in B$, it follows that $x \in A$. Thus, $x \in C$ implies $x \in A$. Therefore, $A \subseteq C$.

(c) Claim: $A \subseteq B$ iff $A \cap B = A$.

"Proof:" Assume that $A \subseteq B$. Suppose $x \in A \cap B$. Then $x \in A$ and $x \in B$, so $x \in A$. This shows that $A \cap B = A$. Now assume that $A \cap B = A$. Suppose $x \in A$. Then $x \in A \cap B$, since $A = A \cap B$; and, therefore, $x \in B$. This shows that $x \in A$ implies $x \in B$, and so $A \subseteq B$.

(d) Claim. For all $n \in Z$, $n \geq 0$, $\frac{(n^2+n)}{2} = \frac{n+(n^2-n)}{2}$.

"Proof:" Consider a set of $n+1$ elements, and let one of these elements be x. There are $\binom{n+1}{n-1} = \frac{n^2+n}{2}$ ways to choose $n-1$ of these elements. Of these, there are $\binom{n}{n-1} = n$ ways to choose $n-1$ elements without choosing x, and $\binom{n}{n-2} = \frac{n^2-n}{2}$ ways to choose $n-1$ elements including x. Therefore, $\frac{n^2+n}{2} = n + \frac{n^2-n}{2}$.

(e) Claim. If the relation R is symmetric and transitive, it is also reflexive.

"Proof:" Since R is symmetric, if $(x, y) \in R$, then $(y, x) \in R$. Thus $(x, y) \in R$ and $(y, x) \in R$, and since R is transitive, $(x, x) \in R$. Therefore R is reflexive.

(f) Claim. If A is a partition of a set A and B is a partition of a set B, then $A \cup B$ is a partition of $A \cup B$.

"Proof:"

i. If $X \in \mathcal{A} \cup \mathcal{B}$, then $X \in \mathcal{A}$, or $X \in \mathcal{B}$. In either case, $X \neq \emptyset$.

ii. If $X \in \mathcal{A} \cup \mathcal{B}$, and $Y \in \mathcal{A} \cup \mathcal{B}$, then $X \in \mathcal{A}$ and $Y \in \mathcal{A}$, or $X \in \mathcal{A}$ and $Y \in \mathcal{B}$, or $X \in \mathcal{B}$ and $Y \in \mathcal{A}$, or $X \in \mathcal{B}$ and $Y \in \mathcal{B}$. Since both $\mathcal{A}$ and $\mathcal{B}$ are partitions, in each case either $X = Y$ or $X \cap Y = \emptyset$.

iii. Since $\bigcup_{x \in \mathcal{A}} X = A$ and $\bigcup_{x \in \mathcal{B}} X = B$, $\bigcup_{X \in \mathcal{A} \cup \mathcal{B}} X = A \cup B$.

2. If A, B, C and D are sets, then $(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$.

3. Let $\mathcal{A} = \{A_\alpha : \alpha \in \Delta\}$ be a family of sets, $\Delta \neq \emptyset$, and B be a set. For each

statement either prove the statement is true or give a counterexample.

(a)
$$B - \left(\bigcap_{\alpha \in \Delta} A_\alpha \right) = \bigcap_{\alpha \in \Delta} (B - A_\alpha)$$

(b)
$$\left(\bigcup_{\alpha \in \Delta} A_\alpha \right) - B = \bigcup_{\alpha \in \Delta} (A_\alpha - B)$$

4. The number of four-digit numbers that can be formed using exactly the digits $1, 3, 3, 7$ is less than $4!$, because the two 3's are indistinguishable. Prove that the number of permutations of n objects, m of which are alike, is $\frac{n!}{m!}$. Generalize to the case when m_1 are alike and m_2 others are alike.

5. Let R be a relation on a set A that is reflexive and symmetric, but not transitive. Let $R(x) = \{y : xRy\}$. Does the set $\mathcal{A} = \{R(x) : x \in A\}$ always form a partition of A? Prove your answer is correct.

6. Prove that $x = y \pmod{m}$ is an equivalence relation on $\mathbb{Z}$.

Reflection Questions

After completing the task, students were asked to reflect on their work with the following questions.

1. Will you describe any strategies that you used while working on this task?

2. Why did you choose this strategy?

3. If the student attempted to construct an example:

   (a) For what purpose did you construct an example?

   (b) How did you construct this particular example? Why did you choose these numbers, etc.?

(c) Did you start with the hypotheses or the conclusion?

(d) Do you think your construction is an example of the definition?

4. Do you think you implemented the strategy effectively? Explain.

5. What other strategies could you have used?

APPENDIX F

THIRD INTERVIEW PROTOCOL WITH THE STUDENTS

Initial Questions

These questions were asked before giving the students' the tasks.

1. How do you use examples in this class?
2. Can you think of an instance (or more than one) where the instructor used an example that was helpful?
3. For what purposes to you consider examples?
4. What was the purpose of the examples that the instructor used that you found helpful?
5. What do you learn from constructing an example?
6. What did you learn from the examples constructed by the instructor?
7. How do you construct examples? What process do you use?

Tasks

The students were given approximately 45 to 60 minutes to work on the following problems. They were asked to show all scratch work, and to think out loud.

1. Prove or give a counterexample:
   (a) If f and g are decreasing functions on an interval I, and $f \circ g$ is defined on I, then $f \circ g$ is decreasing on I.

(b) If f and g are decreasing functions on an interval I, and $f \circ g$ is defined on I, then $f \circ g$ is increasing on I.

2. Suppose f is an increasing function. Prove or disprove that there is no real number c that is a global maximum for f.

A real-valued function is called *fine* if it has a root (zero) at each integer.

3. Give an example of a fine function and explain why it is a fine function.

4. Prove or give a counterexample:

   (a) All fine functions are periodic. (Hint: A real-valued function f is periodic if there exists a number $c \in \mathbb{R}$ such that $f(x+c) = f(x)$ for all $x \in \mathbb{R}$).

   (b) The product of a fine function and any other function is a fine function.

   (c) Let $f : \mathbb{R} \to \mathbb{R}$ be a fine function. Let $g : \mathbb{R} \to \mathbb{R}$ be a function defined by $g(x) = f(x-k)$ for some $k \in \mathbb{Z}$. Prove that g is a fine function.

5. Let $f : \mathbb{R} \to \mathbb{R}$ be a fine function. If $g : \mathbb{R} \to \mathbb{R}$ is a function, such that $g \circ f$ is a fine function, what conditions must g satisfy? Does g have to be a fine function as well? What is the weakest condition that g must satisfy to ensure the composition is fine?

Note: Question 2 is from Alcock and Weber (2010). Question 3-5 were inspired by Dahlberg and Housman (1997) and Iannone et al. (2011).

Reflection Questions

After completing the task, students were asked to reflect on their work with the following questions.

1. If the student attempted to construct an example:

(a) For what purpose did you construct an example?

(b) How did you construct this particular example? Why did you choose these numbers, etc.?

(c) Did you start with the hypotheses or the conclusion?

(d) Do you think your construction is an example of the definition?

2. Do you think you implemented the strategy effectively? Explain.

APPENDIX G

INTERVIEW PROTOCOL WITH THE INSTRUCTOR

These are some of the questions that will be asked of the instructor.

1. Why did you decide to include [insert specific example] in your lecture today?
2. What do you want students to be able to do based on [the example] you chose today?
3. What is your opinion on the work that [insert student's name] did today? Do you approve of the approach she took?

REFERENCES

Alcock, L. (2009). Teaching proof to undergraduates: semantic and syntactic approaches. In F. Lin, F. Hanna, & M. de Villiers (Eds.), *Proceedings of the icmi study 19 conference: Proof and proving.* (Vol. 1, p. 227-246). Taipei: National Taiwan Normal University.

Alcock, L. (2010). Mathematicians' perspectives on the teaching and learning of proof'. In F. Hitt, D. Holton, & P. W. Thompson (Eds.), *Cbms numbers in mathematics education, research in collegiate mathematics education vii* (Vol. 16, p. 63-92).

Alcock, L., & Inglis, M. (2008). Doctoral students' use of examples in evaluating and proving conjectures. *Educational Studies in Mathematics*(2), 111-129.

Alcock, L., & Inglis, M. (2009). Representation systems and undergraduate proof production: A comment on Weber. *Journal of Mathematical Behavior*(28), 209-211.

Alcock, L., & Simpson, A. (2004). Convergence of sequences and series: Interactions between visual reasoning and the learner's belief about their own role. *Educational studies in mathematics*, *57*(1), 1-32.

Alcock, L., & Simpson, A. (2005). Convergence of sequences and series 2: Interactions between nonvisual reasoning and the learner's belief about their own role. *Educational studies in mathematics*, *58*(1), 77-100.

Alcock, L., & Weber, K. (2005). Proof validation in real analysis: Inferring and checking warrants. *Journal of Mathematical Behavior*, *24*, 125-134. doi: 10.1016/j.jmathb.2005.03.003

Alcock, L., & Weber, K. (2009). Referential and syntactic approaches to proving: Case studies from a transition-to-proof course. *Research in Collegiate Mathematics Education*, *7*, 101-123.

Alcock, L., & Weber, K. (2010). Undergraduates' example use in proof construction: Purposes and effectiveness. *Investigations in Mathematics Learning*, *3*(1), 1-22.

Antonini, S. (2006). Graduate students' processes in generating examples of mathematical objects. In *Proceedings of the 30th international conference on the psychology of mathematics education* (Vol. 2, p. 57-64). Prague, Czech Republic.

Bills, L., & Tall, D. (1998). Operable definitions in advanced mathematics: The case of the least upper bound. In A. Olivier & K. Newstead (Eds.), *Proceedings of the twenty-second international conference for the psychology of mathematics education* (Vol. 2, p. 104-111). Stellenbosch, South Africa: University of Stellenbosch.

Buchbinder, O., & Zaslavsky, O. (2011). Is this a coincidence? the role of examples in fostering a need for proof. *ZDM Mathematics Education*, *43*, 269-281. doi: 10.1007/s11858-011-0324-7

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

Charmaz, K. (2006). *Constructing grounded theory.* Thousand Oaks, CA: Sage Publications Ltd.

Clarke, D., Breed, M., & Fraser, S. (2004). The consequences of a problem-based mathematics curriculum. *The Mathematics Educator*, *14*(2), 7-16.

Cobb, P., Stephan, M., & Gravemeijer, K. (2001). Participating in classroom mathematical practices. *Journal of the Learning Sciences*, *10*(1-2), 113-163.

Creswell, J. W. (2013). *Qualitative inquiry and research design* (3rd ed.). Thousand Oaks, CA: Sage.

Dahlberg, R. P., & Housman, D. L. (1997). Facilitating learning events through example generation. *Educational Studies in Mathematics*, *33*(3), 283-299.

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

Dreyfus, T. (1991). Advanced mathematical thinking. In P. Nesher & J. Kilpatrick (Eds.), *Mathematics and cognition: A research synthesis by the international group for the psychology of mathematics education.* Cambridge: Cambridge University Press.

Epp, S. S. (2003). The role of logic in teaching proof. *The American Mathematical Monthly*, *110*(10), 886-899.

Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. Resnick (Ed.), *The nature of intelligence* (p. 231-236). Hillsdale, NJ: Erlbaum.

Fukawa-Connelly, T. P. (2012a). A case study of one instructor's lecture-based teaching of proof in abstract algebra: Making sense of her pedagogical moves. *Educational Studies in Mathematics*, *81*, 325-345.

Fukawa-Connelly, T. P. (2012b). Classroom sociomathematical norms for proof presentation in undergraduate in abstract algebra. *Journal of Mathematical Behavior*, *31*, 401-416.

Garofalo, J., & Lester, F. K., Jr. (1985). Metacognition, cognitive monitoring, and mathematical performance. *Journal for Research in Mathematics Education*, *16*(3), 163-176.

Glaser, B. G., & Strauss, A. (1967). *The discovery of grounded theory.* Chicago: Aldine.

Glesne, C. (2011). *Becoming qualitative researchers* (4th ed.). Boston, MA: Pearson.

Goldin, G. A. (1998). Representational systems, learning, and problem solving in mathematics. *Journal fo Mathematical Behavior*, *17*(2), 137-165.

Harel, G., & Sowder, L. (2007). Toward comprehensive perspectives on the learning and teaching of proof. In F. K. Lester (Ed.), (p. 805-842). Charlotte, NC: Information Age.

Harel, G., & Tall, D. (1991). The general, the abstract, and the generic in advanced mathematics. *For the Learning of Mathematics*, *11*(1), 38-42.

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

Hodds, M., Alcock, L., & Inglis, M. (2014). Self-explanation training improves proof comprehension. *Journal for Research in Mathematics Education*, *45*, 62-101.

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

Inglis, M., & Alcock, L. (2012). Expert and novice approaches to reading mathematical proofs'. *Journal for Research in Mathematics Education*, *43*(4), 358-390.

Inglis, M., & Alcock, L. (2013). Skimming: A response to weber and mejia-ramos. *Journal for Research in Mathematics Education*, *44*(2), 472-474.

Knuth, E. (2002). Secondary school mathematics teachers' conceptions of proof. *Journal for Research in Mathematics Education*, *43*, 358-390.

Lakatos, I. (1976). *Proofs and refutations.* Cambridge: Cambridge University Press.

Larsen, S., & Zandieh, M. (2008). Proofs and refutations in the undergraduate mathematics classroom. *Educational Studies in Mathematics*, *67*, 205-216. doi: 10.1007/s10649-9106-0

Laursen, S. L., Hassi, M.-L., Kogan, M., & Weston, T. J. (2014). Benefits for women and men of inquiry-based learning in college mathematics: A multi-institution study. *Journal for Research in Mathematics Education*, *45*, 406-418.

Leron, U., & Dubinksy, E. (1995). An abstract algebra story. *American Mathematical Monthly*, *90*(3), 174-184.

Lester, F. K., Jr. (1994). Musings about mathematical problem-solving research: 1970-1994. *Journal for Research in Mathematics Education*, *25*(6), 660-675.

Lester, F. K., Jr., & Garofalo, J. (1982, March). *Metacognitive aspects of elementary school students' performance on arithmetic tasks.* Paper presented at the metting of the American Educational Research Association, New York.

Lew, K., Fukawa-Connelly, T. P., Mejia-Ramos, J. P., & Weber, K. (2015). Lectures in advanced mathematics: Why students might not understand what the mathematics professor is trying to convey. *preprint*.

Lew, K., Mejia-Ramos, J. P., & Weber, K. (2013). Not all informal representations are created equal. In S. Brown, K. H. Roh, & M. Oerhtman (Eds.), *Proceedings of the 16th annual conference on research in undergraduate mathematics education* (p. 570-574). Denver, CO.

Lockwood, E., Ellis, A., & Knuth, E. (2013). Mathematicians example-related activity when proving conjectures. In S. Brown, K. H. Roh, & M. Oerhtman (Eds.), *Proceedings of the 16th annual conference on research in undergraduate mathematics education* (p. 21-23). Denver, CO.

Lockwood, E., Ellis, A. B., Dogan, M. F., Williams, C. C., & Knuth, E. (2012). A framework for mathematicians' example-related activity when exploring and proving mathematical conjectures. In *Proceedings of the 34th annual meeting of the north american chapter of the international group for the psychology of mathematics education* (p. 151-158).

Mason, J., & Pimm, D. (1984). Generic examples: Seeing the general in the particular. *Educational Studies in Mathematics*, *15*, 227-289.

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

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

Mejia-Ramos, J. P., & Weber, K. (2015). Mathematics majors' example and diagram usage when writing calculus proofs. In *Proceedings of the 17th conference for research in undergraduate mathematics education.*

Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation.* San Francisco, CA: Jossey-Bass.

Michener, E. R. (1978). Understanding understanding mathematics. *Cognitive Science*, *2*, 361-383.

Mills, M. (2014). A framework for example usage in proof presentations. *The Journal of Mathematical Behavior*, *33*, 106-118.

Mills, M., Brown, S., Larsen, S., Marrongelle, K., & Oehrtman, M. (2011). Mathematicians pedagogical thoughts and practices in proof presentation. In *Proceedings of the 14th conference for research in undergraduate mathematics education* (Vol. 2, p. 283-297).

Moore, R. C. (1994). Making the transition to formal proof. *Educational Studies in Mathematics*, *27*(3), 249-266.

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

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.

Pedemonte, B., & Buchbinder, O. (2011). Examining the role of examples in proving processes through a cognitive lens: the case of triangular numbers. *ZDM Mathematics Education*, *43*, 257-267. doi: 10.1007/s11858-011-0311-z

Pinto, M., & Tall, D. (2002). Building formal mathematics on visual imagery: A case study and a theory. *For the Learning of Mathematics*(1), 2-10.

Polya, G. (1957). *How to solve it: a new aspect of mathematical method* (2nd ed.). Princeton, NJ: Princeton University Press.

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

Rasmussen, C., & Stephan, M. (2008). A methodology for documenting collective activity. In A. E. Kelly & R. Lesh (Eds.), *Design research in education.* Mawah, NJ: Erlbaum.

Rowland, T. (2001). Generic proofs in number theory. In S. R. Campbell & R. Zazkis (Eds.), *Learning and teaching number theory: Research in cognition and instruction* (p. 157-183). Westport, CT: Albex Publishing.

Sandefur, J., Mason, J., Stylianides, G., & Watson, A. (2012). Generating and using examples in the proving process. *Educational Studies in Mathematics*, 1-18.

Schoenfeld, A. H. (1980). Teaching problem-solving skills. *The American Mathematical Monthly*, *87*(10), 794-805.

Schoenfeld, A. H. (1981, April). *Episodes and executive decisions in mathematical problem solving.* Paper presented at the meeting of the American Educational Research Association, Los Angeles.

Schoenfeld, A. H. (1983). Beyond the purely cognitive: Belief systems, social cognitions and metacognitions as driving forces in intellectual performance. *Cognitive Science*, *7*, 329-363.

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense-making in mathematics. In D. Grouws (Ed.), *Handbook for research on mathematics teaching and learning* (p. 334-370). New York: MacMillan.

Selden, J., & Selden, A. (1995). Unpacking the logic of mathematical statements. *Educational Studies in Mathematics*, *29*(2), 123-151.

Seymour, E., & Hewitt, N. (1997). *Talking about leaving: Why undergraduates leave the sciences.* Boulder, CO: Westview Press.

Smith, J. C. (2006). A sense-making approach to proof: Strategies of students in traditional and problem-based number theory courses. *Journal of Mathematical Behavior*, *25*, 73-90.

Sowder, L., & Harel, G. (2003). Case studies of mathematics majors' proof understanding, production, and appreciation. *Canadian Journal of Science, Mathematics & Technology Education*, *3*(2), 251-267.

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

Suritsky, S. K., & Hughes, C. A. (1991). Benefits of notetaking: Implications for secondary and postsecondary students with learning disabilities. *Learning Disability Quarterly*, *14*(1), 7-18.

Texas State University-San Marcos. (2013). *Undergraduate catalog, 2012-2014*. Retrieved from `http://www.txstate.edu/curriculumservices/catalogs/undergraduate/catalogs/2012-14.html`

Vinner, S. (1991). The role of definitions in the teaching and learning mathematics. In D. O. Tall (Ed.), *Advanced mathematical thinking* (p. 65-81). Netherlands: Springer.

Watson, A., & Mason, J. H. (2002). Extending example spaces as a learning strategy in mathematics. In A. Cockburn & E. Nardi (Eds.), (Vol. 4, p. 377-384). Norwich, UK: University of East Anglia.

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

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

Weber, K. (2002). Beyond proving and explaining: Proofs that justify the use of definitions and axiomatic structures and proofs that illustrate technique. *For the Learning of Mathematics*, *22*(3), 14-17.

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

Weber, K. (2005a). Problem-solving, proving, and learning: The relationship between problem-solving processes and learning opportunities in the activity of proof construction. *Journal of Mathematical Behavior*, *24*, 351-360. doi: 10.1016/j.jmathb.2005.09.005

Weber, K. (2005b). A procedural route toward understanding aspects of proof: Case studies from real analysis. *Canadian Journal of Science, Mathematics, and Technology Education*, *5*(4), 469-483. doi: 10.1080/14926150509556676

Weber, K. (2008). How mathematicians determine if an argument is a valid proof. *Journal for Research in Mathematics Education*, *39*, 431-459.

Weber, K. (2009). How syntactic reasoners can develop understanding, evaluate conjectures, and generate counterexamples in advanced mathematics. *Journal of Mathematical Behavior*, *28*, 200-208.

Weber, K. (2010). Proofs that develop insight. *For the Learning of Mathematics*, *30*(1), 32-36.

Weber, K. (2012). Mathematicians' perspectives on their pedagogical practice with respect to proof. *International Journal of Mathematics Education in Science and Technology*, *43*, 463-482.

Weber, K. (2013). On the sophistication of náive empirical reasoning: factors influencing mathematicians' persuasion ratings of empirical arguments. *Research in Mathematics Education*, *15*(2), 100-114. doi: 10.1080/14794802.2013.797743

Weber, K., & Alcock, L. (2004). Semantic and syntactic proof productions. *Educational Studies in Mathematics*, *56*(2), 209-234.

Weber, K., & Alcock, L. (2005). Using warranted implication to understand and validate proofs. *For the Learning of Mathematics*, *25*(1), 34-38,51.

Weber, K., Inglis, M., & Mejia-Ramos, J. P. (2014). How mathematicians obtain conviction: Implications for mathematics instruction and research on epistemic cognition. *Educational Psychologist*, *49*(1), 36-58. doi: 10.1080/00461520.2013.865527

Weber, K., & Mejia-Ramos, J. P. (2011). Why and how mathematicians read proofs: an exploratory study. *Educational Studies in Mathematics*, *76*, 329-344.

Weber, K., & Mejia-Ramos, J. P. (2013). On mathematicians' proof skimming: A reply to Inglis and Alcock. *Journal for Research in Mathematics Education*, *44*(2), 464-471.

Weber, K., Porter, M., & Housman, D. (2008). Worked examples and concept example usage in understanding mathematical concepts and proofs. In M. P. Carlson & C. Rasmussen (Eds.), *Making the connection: Research and teaching in undergraduate mathematics education* (p. 245-252). Washington, DC: Mathematical Association of America.

Williams, R. L., & Eggert, A. C. (2002). Notetaking in college classes: Student patterns and instructional strategies. *JGE: The Journal of General Education*, *51*(3), 173-199.