Instructor Perceptions of Student Example Use and Understanding in an Introduction to Proof Course

By

Jordan E. Kirby

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Mathematics Education

Middle Tennessee State University

July, 2023

Dissertation Committee:

Dr. Sarah K. Bleiler-Baxter, Chair

Dr. Alyson Lischka

Dr. Grant Gardner

Dr. James Hart

Dr. Kimberly Evert

ABSTRACT

As students transition from the calculus sequence to upper-level mathematics classes, many struggle to understand the purpose of proofs and how to produce them on their own. One tool students and mathematicians alike frequently use is examples to aid their understanding and production of proof. Using examples productively has been argued to be beneficial to the development of deductive reasoning used in proving mathematical claims. Although much research exists on how students use examples and why they choose the examples to be used, there is little research on how to help instructors aid their students in the productive use of examples. This study seeks to answer this lack of research by asking the research question: How do instructors perceive the example-use of their students in an introduction to proof course?

Instructors of introduction to proof courses from across the southeastern United States were interviewed and shown samples of student work with varying degrees of effectiveness in using examples to answer two different questions. Participants of the interview later categorized the sample student work and finally ranked the student responses for how well they believed the students gave a proof of the problem presented. Qualitative analysis of the interviews indicated instructors commonly ranked and categorized the student responses in alignment with mathematics education research related to example-use. Although most participants recommended students perceived as struggling with the task use examples to help build intuition to solve the problems, participants were split on whether students commonly ranked at the top should include examples in the formal written work presented in class.

Other findings indicated participants typically suggested their students use examples along a continuum of example-use rather than jump straight to formal proof excluding examples. Common instructional responses to students at varying levels of example-use were again aligned with mathematics education research and aimed to help students advance in understanding along a continuum of responses. Finally, a lesson was developed to implement the research findings discussed in the first part of this dissertation in an introduction to proof course with the goal of highlighting the benefits of examples to both students and instructors.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Introduction to Proof (ITP)

Journal of Mathematical Behavior (JMB)

Mathematical Knowledge for Teaching (MKT)

Mathematical Knowledge for Teaching proof (MKT-P)

Pedagogical Content Knowledge (PCK)

Subject Matter Knowledge (SMK)

Knowledge of Content and Students (KCS)

Knowledge of Content and Teaching (KCT)

Chapter 1 Introduction

Mathematics policy documents in the 21st century have argued that a strong understanding of proof should be positioned as a key educational goal for students of mathematics (National Governors Association (NGA)/ Council of Chief State School Officers, 2010; RAND Mathematics Study Panel, 2002). *Principles and Standards* (National Council of Teachers of Mathematics (NCTM), 2000) and the *Common Core State Standards Initiative for Mathematics* (NGA/CCSSI, 2010) have continued to emphasize a goal in mathematics to have students develop deductive reasoning and argumentation aided through proof. At the undergraduate level, the Committee on the Undergraduate Program in Mathematics (CUPM) presented a curriculum guide for mathematics majors with an increased emphasis on understanding and producing proofs of substance (CUPM, 2015). This emphasis on proof in K-16 settings has increased efforts of researchers to understand how students come to learn the proving process. These researchers have found students' understanding of proof is lacking compared to the goals set forth by these documents (e.g., Harel & Sowder, 2007; Healy & Hoyles, 2000; Stylianides, 2007; Stylianides & Stylianides, 2017).

A commonly cited difficulty with learning to prove in the K-12 environment comes through the seemingly stark contrast between the introduction to deductive thinking through proof in high school compared to a procedural understanding of mathematics from K-8 (Buchbinder & McCrone, 2020; Marrades & Gutierrez, 2000; Sowder & Harel, 1998), and with this comes a shift from empirical (example-based) argumentation to deductive argumentation when proving. This contrast is mimicked in the undergraduate setting as mathematics majors transition from the calculus sequence to more theoretically based courses, beginning with an introduction to proof (ITP) course. An ITP course provides a setting where students are expected to be producers of their own mathematical ideas for perhaps the first time (Boyle et al., 2015; Harel & Sowder, 2007), and where they are expected to move beyond arguments based on individual examples to arguments that demonstrate the validity of a claim more broadly.

Still, examples in proving and justification in mathematics can aid both proof comprehension and production (Epp, 2003; Selden & Selden, 2003). I am interested in better understanding how examples play a role in the undergraduate transition from calculus-based classes to proof-based classes within ITP courses. A complete definition of how I am using examples as well as instances to better illustrate this definition are given in the "definitions" section below. One quick way to think of how to use an example to aid a proof is to draw a picture or use concrete numbers for an abstract mathematical topic and try to determine what intuition you can gain about the mathematical task you are asked to prove. The purpose of this study is to investigate how instructors perceive their students to use and understand examples in their classroom to aid their understanding and production of proof. This purpose answers a direct call for research from Zaslavsky and Knuth (2019) in the special issue on examples in proving in the *Journal of Mathematical Behavior (JMB):*

> More research is needed on how to elevate the role of proof in ordinary classrooms and how to support teachers' work in enhancing students' understanding of proof. We thus argue that more intervention-oriented studies in the area of proof are sorely needed. (p. 237).

Helping students learn how to prove is a longstanding goal of mathematics education as a field. Studying how students use examples to aid their proving has been one prolific route to understand some of the roadblocks students face when presented with a mathematical task to prove (e.g., Balacheff, 1987; Ellis et al., 2019; Stylianides & Stylianides, 2017). Understanding how the instructors perceive their students to use examples is a necessary step to help prepare instructional materials for practitioners of ITP courses. Although the study proposed here is not an intervention-based study, it lays the foundation for future intervention studies by exploring the baseline perceptions of instructors with respect to students' example-use in proving.

An open and vital question in the example-use in proof literature remains to determine how best to help teachers aid students' understanding of proof (Zaslavsky & Knuth, 2019). Most research in the example-use in proof literature has focused on how students understand and use examples when proving a mathematical claim (Aricha-Metzer & Zaslavsky, 2019; Lynch & Lockwood, 2019; Zaslavsky & Knuth, 2019). This study instead focuses on the perceptions of example-use in proving instructors hold of their students in proof classes and what methods instructors use to aid students in achieving a richer understanding of proof. This study will contribute to the field by identifying how instructors currently believe their students to understand example-use in proof. I posit that before a study can intervene to aid instructors in delivering material to their students to help their understanding of example-use, the field must first understand what beliefs instructors hold about their students. Interventions that do not address directly the commonly reported issues instructors believe are held by students will not be as effective. This aspect of the study will help lay the groundwork for studies to answer the call to research from Zaslavsky and Knuth (2019) to begin developing instructional practices to aid students in how to be deliberate with their example-use in a proof setting.

In Table 1 below, I briefly summarize the current state of the field of example-use in proof and how the current study aims to make substantive contributions to the field. In the middle column of Table 1, I give broad questions that this dissertation will seek to make progress in answering for the field. In the far-right column of Table 1, I list some potential implications and future work that can build upon this dissertation to advance the field forward.

Table 1

State of the Literature Related to My Dissertation

What we know	What this study will help us learn	Future directions
· Students struggle with proof (e.g., Harel & Sowder, 2007; Inglis & Alcock, 2012; Stylianides, 2007) · A narrow view of examples by students contributes to a lack of understanding how to prove (e.g., Balacheff, 1988; Buchbinder & Zaslavsky, 2019; Healy & Hoyles, 2000; Zaslavsky et al., 2012) · Through the CAPS framework we know more about HOW students use examples to aid their proving (Ellis et al., 2019) · Generic examples can help students gain insight into how to prove claims (Aricha-Metzer, 2019; Balacheff, 1987)	· What do instructors perceive their students to understand about example-use in proof? · How do instructors perceive their students implement examples in their proof production? · What role do instructors place on their students having an understanding of example-use in proof? · What instructional strategies do instructors suggest for students at varying levels of understanding in proof?	· What instructional methods help students understand how to strategically use examples in a proof setting? · What types of interventions for instructors can be productive in helping these instructors gain more comfortability teaching example-use in proof? · What curriculum or activities can aid instructors in developing rich student understanding of example-use in proof?

Definitions

Definition of example

I define an *example* as taken from Zodik and Zaslavsky (2008) to be "a particular case of a larger class, from which one can reason and generalize." (p. 165). Examples in this sense are used by students to build understanding of a specific scenario to help build conceptual understanding of the more generalizable problem at hand (Iannone et al., 2011; Stylianides & Stylianides, 2009). My research focus, and much of the literature I have read, does not pertain to examples such as problems found in the back of a textbook chapter. The term 'example' has many interpretations and could be considered analogous to exercises for a student to complete. I am instead interested in what is typically used as a step taken by students to help answer a larger question.

Rather than caring about having a student solve for x in some mathematical equation, the literature on example-use is instead interested in asking questions like, why did the student plug in the number 3 in the equation when they got stuck? How did the number 3 specifically help move this student forward? After plugging in the number 3, did the student successfully finish the problem? Would the student have made similar progress if they had instead used the number 5 for this equation? As instructors, how can we help our students know which examples are beneficial to their understanding for a given mathematical problem? The literature on example-use in proof cares about HOW students use the examples (either provided to them or generated by them) to aid in their understanding when proving a mathematical topic. I will define what I mean by proof in the following section.

For an instance of how some researchers investigate examples, Lockwood and colleagues (2016) described an interview where participants were asked about the examples they created to guide their thinking when presented with a new mathematical idea to prove. There was no specific task given. Rather, participants in this study were asked broadly what strategies they would use to try and figure out how to prove a mathematical idea. Responses to this question included participants responding with ideas including “boundary examples” (p. 176) where a participant tried to break the conjecture with the given example. One participant noted,

> I try to “break” the conjecture; that is, I choose examples that are likely to show the conjecture is not true. For number-based conjectures, I choose 0, numbers close to 0 (both positive and negative), very large and very small numbers, for example, both integers and non-integers. (p. 178)

The specific numbers or pictures drawn by participants when using examples are not always the interest of study. The findings described by Lockwood and colleagues were instead interested in how these general strategies would be applicable for many different situations or students. By understanding how some individuals successfully gain intuition to complete a mathematical problem or proof when using examples, we as mathematics education researchers better understand how to help those individuals who are unsuccessful when trying to solve a mathematical problem or proof.

Although examples can be used in many different mathematical settings, I am particularly interested in how examples can aid students in gaining intuition for how to prove a mathematical claim. These examples can look like what has been described above, starting with specific numbers to plug into a formula or a picture to help see the mathematical structure of a problem. One line of research in the field of example-use in proof is to investigate *how* students use examples to aid their proof production and understanding as opposed to *why* students used a specific example in their approach.

Definition of Proof

I draw upon the definition of proof as taken from Stylianides and Stylianides (2017): "A mathematical argument for or against a mathematical claim that is both mathematically sound and conceptually accessible to the members of the local community where the argument is offered." (p. 212). When determining what makes an argument or claim "mathematically sound" as described by Stylianides and Stylianides, I refer to the proof scheme framework developed by Harel and Sowder (2007).

One of the primary frameworks for understanding how students understand proof comes from Harel and Sowder (2007). In their chapter, Harel and Sowder described the different ways students understand verification in a mathematics context. Harel and Sowder (2007) argued students understand verification in mathematics from three different perspectives: *external conviction* proof scheme, *empirical* proof scheme, and *deductive* proof scheme. In an external conviction proof scheme, students were characterized as understanding the validity of an argument based on external factors including the authority of the figure (i.e., a teacher of the class or the words in a textbook), the physical appearance of the book (i.e., two-column proofs), or symbolic manipulations showing some gap of knowledge in the student by breaking a rule of a system. Students holding an *empirical* proof scheme may be convinced of the validity of an argument based upon evidence from an example or initial conceptions of a topic without the mathematical rigor required to effectively argue for validity.

The final categorization described by Harel and Sowder (2007) was a *deductive* proof scheme. The deductive proof scheme is apparent in student reasoning when students make progress in generating or understanding a generic argument through logical inference or logical steps. Students in this outcome form goals about the problem presented to be proven and anticipate the outcomes of their actions.

For more clarity on what is meant by a proof, consider your friend telling you something they think might be true, "I think whenever I add two odd numbers together, the result will always be an even number." We as mathematicians might refer to this statement as a claim or a conjecture. To prove this claim meeting the definition set out by Stylianides and Stylianides (2017), we would be tasked with creating a mathematically sound argument we could then tell our friend and convince them of the validity of our steps. It might be tempting to tell your friend, "Well, I noticed when I added 3 and 5 together that made 8. So, I think you are right!" However, this is not a mathematically sound argument, because a single instance of a claim being true is insufficient to prove that every case of this claim is true. To relate to the proof scheme framework described by Harel and Sowder (2007), this argument might be indicative of a student with an empirical proof scheme. Is this claim made by your friend true for all numbers? Is it true for both positive and negative numbers? What if you add the same number to itself? These are all questions we as mathematicians must ask ourselves when considering steps to prove a claim.

To gain intuition for how to solve this claim, you may use an (abstracted) example. For instance, we can look back at the single instance of numbers used above. Someone exhibiting a deductive proof scheme may instead look to the example of adding

3 and 5 and find some intuition. Three is a group of two with one left over. Five is two groups of two with one left over. If you take these two leftover pieces and add them together, you will get another whole group of two. Having all groups of two with no leftover pieces means you will have an even number. Notice how instead of relying on a single instance being sufficient for proving the case above, this argument starts to get at the underlying structure of what an odd number or an even number means by definition. A proof is an argument like this in typically written form for the purposes of communicating to someone else.

Significance

This study will contribute to the field of example-use in proof in three main ways. The main contribution of this study will be to lay the groundwork to answer the call from the 2019 JMB special edition for how to give support to instructors of ITP classes to teach example-use in a productive manner to students. Zaslavsky and Knuth (2019) note:

> In other words, there is a need to address the voids in both the guidance available to practitioners and the research base upon which such guidance is founded, with respect to teaching and learning to use examples strategically and productively in the course of conjecturing and proving. (p. 243)

To answer this call, this study will investigate the current perceptions instructors of ITP courses hold about their students' understanding and use of examples in proof. These perceptions will guide future research to give advice to practitioners of these courses on the best practices for strategically implementing example-use in proof.

The second contribution of this study will be to organize common instructional practices used by the participants of my study. These participants will be instructors of ITP courses and over the course of a semi-structured interview will be asked how to respond to various student work samples. Some of these student work samples will show a student who is struggling to understand the concepts given in the problem. Other student work samples in this interview will highlight students who give a solid answer to the question they are asked. Participants in this study will respond how they would help both struggling students and students who are showing solid understanding if these students were in the instructors' classes.

Finally, the third contribution of this study will be to present practitioners with a task that can be implemented in an ITP course. After conducting interviews with instructors of ITP courses across the country, I will formalize my interview protocol to be a lesson applicable to an early semester ITP course. I intend this to be shared with instructors as a preliminary step to help answer the calls from Zaslavsky and Knuth (2019) above with the preparation of how to help instructors get their students to use examples more effectively.

Research Questions

An overarching research question for my dissertation is: How do instructors perceive the example-use of their students in an ITP course? To make progress in answering this question, I will introduce two smaller questions and one practitioner article with the intention of forming three publishable manuscripts. This will constitute the alternative format of this dissertation. My first research question (referred to as RQ1) is: How do instructors of ITP classes perceive students' understanding and use of examples when proving? In Manuscript 1 for my alternative format dissertation, I will discuss some of the primary findings of my study related to the perceptions held by participants about their students' use of examples in an ITP course. My second research question (referred to as RQ2) is: What instructional strategies do instructors of ITP courses suggest for students using examples to aid their proving? Exploring and answering this question will constitute Manuscript 2 for my alternative format dissertation. My third manuscript will consist of an article geared toward practitioners with a focus on how to implement a task early in the semester of an ITP course. The goal of this task will be to help develop students' intuition about how to use examples productively to aid their proving. I will conclude the alternative format dissertation by summarizing my findings from the three separate manuscripts and stating implications for practice and future research.

Personal Background and Interest

My personal background includes identities as both a mathematician and a mathematics educator. I have both a bachelor's degree and master's degree from Middle Tennessee State University (MTSU) in mathematics. I am currently finishing a Ph.D. in mathematics education at the same university. In this section, I will briefly share my personal interest in the topic as well as my position as an education researcher.

I do not personally subscribe to the idea of a dream job for everyone. I do not think the necessity of working should be glorified with some idealistic reality of work that satisfies someone. With that being said, teaching has and will always be my dream job. In high school, I was the president of the Future Teachers of America for three years and a member for four. When it came time to declare a major for college, I immediately chose to be a secondary mathematics education major in the MTeach program. Mathematics was always one of the only fields that interested me. Memorization bored me, and the precision required of some sciences scared me. However, mathematics fascinated me. With sufficient understanding, you could recreate needed formulas for yourself and eliminate the need for much of the memorization in mathematics. This line of thinking led me to start tutoring mathematics in the eighth grade and continue for all ten years of college both for MTSU and working for a private tutoring company. In total, I have spent well over 1500 hours of my life tutoring mathematics in a 1-on-1 setting and never had a dull moment. Then I took Foundations of Higher Mathematics, the course serving as an introduction to proofs at MTSU. After this course, I changed my major from the MTeach program to a pure mathematics major to investigate more mathematics.

I never struggled with mathematics growing up. I rarely required time to sit down and study. As I sat in a class, be it algebra or calculus, many of the topics would become clear to me almost immediately and I would be able to recreate what I saw in the classroom. Foundations of Higher Mathematics, however, was a different story. I was only introduced to proof briefly in a ninth-grade geometry classroom where my purpose in proof was to fill out a column on a worksheet and get back to "real" mathematics. As the first proof-based course in my undergraduate mathematics career, Foundations shook me. None of the material covered in this course made sense to me. I struggled every class period to even understand the definitions given to use. I was unable to complete a single homework assignment fully and watched hours of videos trying to understand what was happening. I referred to friends and acquaintances who had taken the class before and started receiving tutoring myself. Still, nothing made sense to me. Then, a lightbulb. In the final week of the course something in my brain clicked. I realized the intuition I had been using in my calculus sequence or algebra classes was now just put into words for a more general sense. I went from almost failing the class to getting a high A on the final exam and overall, a C for the course. I knew that week I was hooked.

The next semester I changed my major to a pure mathematics major. The MTeach program didn't offer enough mathematics courses to satiate my desire to learn more. I changed my goal to instead teach at the collegiate level and maybe even to teach proof-based classes like the one I struggled so much with. In the second year of my master's program, I was able to start teaching college as an instructor of record and knew I made the right decision. Now five years later, I still look forward to waking up every morning and having the opportunity to teach at college.

When I started researching mathematics education, the only field I was interested in was proof research. As I continued to study the various aspects of mathematics education, I noticed how lucky I was to go to a college with as many progressive mathematics instructors as I had. The nature of mathematics and science class I took in this program helped open my eyes to how many people hold a strong almost didactic view of mathematics. This was extended in my STEM integration class where consistently mathematics was just seen as a tool to help do other things rather than a field of study on its own. I decided to research the instructor side of teaching these proof classes and see what I could discover from people holding these wildly different views of mathematics than I do that might exist in my same field.

When I read the Harel and Sowder (2007) chapter as part of my coursework, I knew I was interested in example-use in proof and proof schemes for my dissertation. I thought back to my own personal understanding of proof and how I began my ITP class with what Harel and Sowder would classify as an external proof scheme. Over the course of the semester, my understanding of proof slowly evolved and became less formulaic and focused on memorizing when to use certain words or phrases. Instead, I began to focus on the mathematics at hand and trying to understand the concepts of mathematics like I did in my calculus classes. I saw classmates hold a more external proof scheme or formulaic understanding of my upper division mathematics classes as I continued to study.

It fascinated me that seemingly so many people could hold these views of a subject I love so different from my own. Proof schemes became something of a frame through which I began to wonder how the instruction of ITP classes could help students combat their limited view of the nature of mathematics or conceptual understanding of proof. Example-use in proof became a lens through which I became interested in investigating how different students or mathematicians developed their understanding of proof. Although my study has evolved to lean less on the work of Harel and Sowder's (2007) proof scheme framework, I still credit their chapter with instigating my curiosity for the subject of example-use in proof. I hope to leverage this experience of my own personal identity as a mathematician struggling in a course with that of instructors across the country teaching the same course I struggled in. I want to investigate how they perceive their students to understand this topic of example-use due to its generality and applicability to all upper division mathematics.

Summary

This study aims to investigate the perceptions instructors of ITP classes hold of their students' example-use. Example-use in proof received a special issue of JMB in 2019 highlighting current work in the field and needed future directions. To answer the call in this issue of how to prepare instructors to teach example-use effectively to their students, I will conduct a large-scale study from which I will have three separate manuscripts aiming to answer my overarching research question: How do instructors value the example-use of their students in an ITP course? Answering this question will lay groundwork for the field of example-use in proof research to begin to develop instructional interventions and answer the call from the 2019 JMB article of how to productively teach example-use in an ITP classroom. I will investigate first how instructors perceive their students to use examples.

Article 1: Instructors Perceptions (Intended for Journal of Mathematical Behavior)

How Instructors Perceive Their Students to Use Examples for Proof

1. Introduction

The role of proof and proving in mathematics education research is well researched (e.g., Harel & Sowder, 2007; Inglis & Alcock, 2012; Stylianides, 2007a). A common finding of the research conducted on student struggles with understanding proof is the difficulty in understanding the role of examples to aid proof production and conceptions (e.g., Balacheff, 1988; Buchbinder & Zaslavsky, 2019; Healy & Hoyles, 2000; Zaslavsky et al., 2012). Recent research has helped characterize how students use examples in proving (e.g., Aricha-Metzer & Zaslavsky, 2019; Ellis et al., 2019). As noted by Zaslavsky and Knuth in the 2019 special issue on example-use of the *Journal of Mathematical Behavior,* however, "very little research has focused on the nature and design of instructional practices that facilitate the development of students' abilities to strategically think about and productively use examples as they engage in proving-related activities." (p. 243). This manuscript will help provide information for the current state of practitioners methodologies and perceptions regarding example-use in proof through responses to task-based interviews.

This manuscript lays the groundwork for future research addressing a suggestion posed by Zaslavsky and Knuth (2019), namely, "to design and test instructional interventions intended to help students learn to strategically think about and productively use examples in proving-related activities." (p. 243). By first understanding the beliefs and perceptions of instructors, researchers in the field of example-use can gain insight into how closely practitioners align their beliefs with the recommended strategies used by researchers. This check for alignment will help future researchers design more productive interventions for instructors of introduction to proof (ITP) classes. Instructional interventions will be critical in assisting instructors of ITP courses to help their students learn to prove more effectively.

1.2 Research question

By gaining an understanding of how instructors of ITP courses currently understand their students to be using examples in their classrooms, mathematics education researchers develop more appropriate interventions in the future targeting instructors where they are at, rather than where we hope they will be. To progress this research, this study will analyze how instructors of ITP classes react and respond to student work samples including examples in their proofs. This study answers the following research question in this manuscript:

1. How do instructors of ITP classes perceive students' understanding and use of examples?

2. Theoretical background

The theoretical framing of this study has two parts: (1) definition of proof as a social construct and, (2) examples in proof. First, setting up the definition of proof positions my understanding of proof along with components of proof like formality and community. Describing proof as a social construct is necessary to understand the context in an ITP class. Second, by introducing the various ways that examples are used in proof, how students may use examples to aid their intuition in proof can be further explained. Further, this explanation will ground the interview protocol for this study. These frameworks will allow help to accomplish the goal of this manuscript to understand how instructors of ITP classes perceive students' understanding and use of examples. The context for this study includes student work from an ITP course at the undergraduate level. ITP courses are often the first dedicated class for undergraduates to focus on proof and proving. In the context of my study, all ITP courses required calculus 2 as a prerequisite.

2.1 Definition of proof

There are many definitions of proof in the education literature (e.g., Balacheff, 2002; Reid & Knipping, 2010; Stylianides, 2007b). I draw upon the following definition of proof provided by Stylianides and Stylianides (2017): "A mathematical argument for or against a mathematical claim that is both mathematically sound and conceptually accessible to the members of the local community where the argument is offered." (p. 212). Within this definition, I will focus on the topic of "conceptually accessible."

Proof has been described and researched at varying educational levels. For instance, Stylianides (2007b) investigated how a group of third-grade students could come to understand and prove the claim: "An odd number plus an odd number is an even number." Students in this study were allowed to use everyday language to argue their way through a proof. Stylianides continued to describe his understanding of *continuum,* which asserted, "there should be continuity in how the notion of proof is conceptualized in different grade levels so that students' experiences with proof in school have coherence." (p. 3). Through the lens of his framework, Stylianides determined the student may have an appropriate proof for the elementary school level. To determine if this argument could be considered a proof, the degree to which the classroom community accepted the students' argument was to be considered. This raises an issue of proof as a social construct.

Ernest (1998) described how participants in a social setting for proof must either accept or deny the problems raised to them. Further, Stylianides (2007b) citing De Millo and colleagues (1979/1998) said, "[proofs] increase our confidence in the truth of mathematical statements only after they have been subjected to the social mechanisms of the mathematical community." (p. 275). I relate this back to Stylianides and Stylianides (2017) and the notion of conceptually accessible. I interpret a conceptually accessible proof to be a proof in which the community where it is presented understands and accepts the proof. Although the student in Stylianides' study described above lacked some of the rigor and formality associated with proofs at the undergraduate level, the argument was accessible to the community in which it was presented. Similarly, I assert proof within a course can change over time.

In ITP classes, students are faced with being asked to prove a claim on their own for potentially the first time (Boyle et al., 2015). As such, the community of students within the ITP course will shift their understanding of proof over the course of the semester. This means what might be taken as a proof or required in a proof at the beginning of a semester may not be required as the semester progresses. It is therefore important to remember the community in which the proof is presented when discussing what constitutes a proof or the validity of a proof. Understanding proof through this lens helps position myself as a researcher. I assume that examples, when used productively, can be beneficial for students in an ITP course to include in their work.

2.2 Examples in proof

Balacheff (1987) described three categories of how students use examples: naive empiricism, crucial experiment, and generic example. Table 2 details the categories with their accompanying definitions from Balacheff.

Table 2

Balacheff (1987) Categories of Example-Use

Name of category	Definition (from Balacheff 1987, p. 19-20)	My interpretation
Naive Empiricism	Drawing from the observation of a small number of cases the certainty of the truth of an assertion	Based on a small number of instances, students draw a general conclusion about a topic
Crucial Experiment	A process of validation of an assertion in which the individual explicitly poses the problem of generalization and solves it by betting on the realization of a case that he recognizes as being as unspecific as possible	Based on specific cases, a student draws a general conclusion. Rather than just using random instances, the cases for crucial experiment are thought out more and intended to be more representative
Generic Example	Explanation of reasons for the validity of an assertion by carrying out operations or transformations on an object present not for itself, but as a characteristic representative of a class of individuals	Gaining insight from an example to discuss properties of the class of object. Students find an intuition for a pattern or line of reasoning through an example to argue generically

There are many connections between how example-use is discussed in Table 2, and the proof scheme framework discussed by Harel and Sowder (2007). In that work, Harel and Sowder differentiate between an empirical proof scheme and a deductive proof scheme. When students work from an empirical proof scheme, they are typically convinced about the validity of a claim based on a small number of examples. Alternatively, students with a deductive proof scheme tend toward a generic argument through logical inference or logical steps. Balacheff (1987) gives a way to look in more detail at the type of example used by the student or the manner through which a student is using an example. The generic example level of example-use is the most promising to help students toward a deductive proof scheme. Aricha-Metzer and Zaslavsky (2019) noted, “In this sense, a generic example can be seen as a bridge for students, as it may help them move from empirical views and misleading intuitions to a justifiable conviction *that* and an understanding of *why* a statement is true (or false)” (p. 305). This notion was echoed by Buchbinder and McCrone (2020) citing Harel and Sowder, “The authors further suggest that the goal of mathematics education is to help students advance from argumentation based on empirical evidence and authority towards deductive argumentation.” (p. 2).

A generic example can be considered less of a specific example, and more a manner of using an example through which a general argument can be found. I seek to understand how instructors of ITP courses think about their students’ use of examples in varying levels. This study is structured to investigate how instructors would respond to students using examples at each level of Balacheff’s (1987) example-use framework. The intention was to see how instructors of ITP courses reacted differently between what Aricha-Metzer and Zaslavsky (2019) argued were *productive* and *non-productive* ways of using examples for proof.

In their study, Aricha-Metzer and Zaslavsky (2019) made a distinction between productive and unproductive ways of using examples in proof. Aricha-Metzer and Zaslavsky defined a productive way of using an example in proof as:

> Cases were coded as *Productive* if there were indicators that working with the example(s) led the student to make progress towards the development of a proof (or partial proof); that is, if the example use helped the student gain insights into some aspects of the key aspects of the proof, to provide deductive arguments, and/or to articulate a sound justification for *why* the general case holds (p. 307).

Unproductive ways of using examples in proof were described as:

> Cases were coded as *Non-Productive* if the student used example(s) but did not make any progress towards developing or understanding of a proof or a deductive argument (p. 307).

Aricha-Metzer and Zaslavsky found a strong relationship between using examples generically and making progress in proving a task, thereby being classified as a productive way of using an example. Alternatively, there were no cases in which a non-generic use of examples was classified as productive. Aricha-Metzer thus argued using generic examples can help students learn how to prove.

3. Methodology

3.1 Research Design

This study employs a basic interpretive qualitative research design (Merriam, 2002). To answer the research questions posed in this study, it would not be possible to separate the experiences of the participants in how they learn and teach mathematics with how they perceive students to use examples. To understand these nuances and learn more of their stories, interviews were the best choice to accurately represent the participants. Further, the many colloquial uses of the word example could lead to problems with data collection in the use of a survey or other instrument. This led to the beginning of every interview starting with ensuring participants of the study understood what was meant by an example and how the student work shown to them. This design helps ensure that the student work samples shown to participants as part of the task-based interviews conducted with them stayed focused on the Balacheff (1987) framework ideas of example-use. To analyze these task-based interviews, open coding was conducted to search for themes amongst participants' responses to student work. These themes combined with answers to other parts of the interview provide the data to answer how practitioners of ITP courses perceive their students to use examples when proving.

3.2 Participants

The database of Accredited Postsecondary Institutions and Programs (DAPIP; https://ope.ed.gov/dapip/#/search-results) was used to search for all institutions with an accredited mathematics department in Alabama, Arkansas, Florida, Georgia, Kentucky, Mississippi, Missouri, North Carolina, South Carolina, and Tennessee. After compiling a list of all accredited universities with a mathematics department, only those universities with a main campus located in the state of choice were kept. For instance, although Cabrini University is an accredited university with a mathematics department with a campus in Alabama, since their main campus is in Pennsylvania, they were excluded from my data collection.

Further, only universities with the classification of "Institution" on the DAPIP advanced search website were included. No "Additional Location" universities were included. I then used the Carnegie classifications website to denote the classification of each university. After identifying all universities, I went to each university website and found the name and email address of the department chair for mathematics when possible, or the dean of the college in which mathematics was housed when no department chair could be located. I then sent recruitment emails to all these individuals for participation in my study. The southeastern region was chosen to be a sample due to the variation of Carnegie classifications of institutions in the region. Participants in this study are faculty who have recently taught an ITP course across the southeastern United States.

My goal in collecting data was to have a sample close to evenly split between mathematics and mathematics education PhDs. To be included in my sample, instructors must have taught an ITP class in the past, with preference given to those instructors who had taught an ITP class in the past year. Further, I hoped to have a variety of Carnegie classification types of universities represented and a mix of both public and private universities with professors from each category. I also wanted to have multiple professors who had approximately 5 years teaching experience or less, and multiple professors with greater than 15 years of teaching experience.

Finally, I hoped to have multiple professors with a self-described teaching style of "lecture" and multiple professors with a self-described teaching style of "active learning." I use the term actively learning here very broadly and include a range of instructional styles described including constant group work and presentations to activities like think-pair-share. I continued to send emails to participants until I had collected enough participants meeting the above criteria. In total, 259 emails were sent asking for participants of the study. Nineteen professors sent a response email agreeing to participate in my study. I narrowed this pool down to eleven to focus on having a varying pool of the demographics shown in Table 3. These demographics are described in more detail in the following section.

Eleven instructors of ITP classes across the Southeastern United States agreed to participate in the study. Table 3 details the characteristics of my participants. The participants came from six different states and ten different universities. All but one participant had taught introduction to proof within the last two years with some participants noting a different teaching load due to the Covid pandemic.

Table 3

Participant Descriptions

Participant	Carnegie classification	PhD received	Length of time at university	Private or public university	Self-described teaching style
Dr. Sarah	R1	Mathematics Education	5 Years	Public	Active Learning
Dr. Hubert	R1	Mathematics	30 Years	Public	Lecture
Dr. Phil	R2	Mathematics*	18 Years	Public	Lecture
Dr. Ryan	D/PU	Mathematics Education	20 Years	Private	Lecture
Dr. Tucker	M1	Mathematics	7 Years	Public	Active Learning
Dr. Chris	M1	Mathematics Education	18 Years	Public	Lecture
Dr. Daniel	M1	Mathematics	9 Years	Public	Active Learning
Dr. Nathan	Baccalaureate	Mathematics Education	4 Years	Private	Lecture
Dr. Amanda	Baccalaureate	Mathematics*	6 Years	Private	Active Learning
Dr. Casey	Baccalaureate	Mathematics	19 Years	Private	Lecture
Dr. Joe	Baccalaureate	Mathematics Education	1 Year	Public	Active Learning

Participants denoted with a * under PhD received indicated they have publications in mathematics education

The demographics chosen for this study include Carnegie classification of the university, type of Ph.D. received by the interviewee, length of time at the university, whether the university was public or private, and the self-described teaching style used by the interviewee. Pascarella and colleagues (2004) and Seifert and colleagues (2010) argued some liberal arts colleges contribute better to good practices in undergraduate education including active learning, feedback to students, and cooperation among students (Chickering & Gamson, 1987). To ensure as varied responses as possible amongst many possible faculty-student relationships, I wanted a sample with several participants from each cluster of classifications as detailed by Carnegie (carnegieclassifications.acenet.edu). These clusters include baccalaureate colleges, masters' colleges, and doctoral universities. Four participants came from doctoral universities (D/PU, R1, R2), three participants came from masters universities (M1,M2,M3), and four participants came from baccalaureate universities.

Although it is known that experience can impact student achievement and understanding in mathematics (e.g., Cho & Baek, 2019; Clotfelter et al., 2007), there is less research at the postsecondary level on how the type of Ph.D. received (i.e., mathematics versus mathematics education) affects student achievement. I assume in this study that the type of Ph.D. received may affect the participants familiarity with frameworks like Balacheff's (1987) framework described above and therefore affect how the participants perceive their students to use examples. For clarity, no participants of this study when asked were familiar with Balacheff's framework or the different ways to use examples in proof. There were five participants in this study with a mathematics education Ph.D., and six participants with a mathematics Ph.D. Two of the participants with a mathematics Ph.D. had published at least one article in a mathematics education journal and have an asterisk in Table 3.

Cho and Baek (2019) argued rank of professor and length of time teaching at the university level affects the quality of teaching in science and mathematics classrooms. To ensure a variety of ranges for my sample, I collected data until I had a population of at least one full professor and one associate professor at each Carnegie classification rank. Four participants of this study were at the assistant professor rank, two participants were the rank of associate professor, and five participants were at the rank of full professor. In searching for as varied a sample as possible, I sought to include professors in both private and public university settings. I am unfamiliar with any work currently investigating the relationship between private or public universities and teaching styles, however I thought to include the distinction in case something came out in the results. Four participants in this study were at private universities and seven participants were at public universities.

Finally, I made sure to include at least one participant who self-identified using a teaching style of lecture or a traditional style and one participant who self-identified using an active learning teaching style in each cluster of Carnegie classifications. Instructor pedagogy is known to affect students' attitude toward mathematics (Sonnert et al., 2014) as well as be a contributing factor to teaching effectiveness as perceived by students (Freeman et al., 2014; Opdenakker & Van Damme, 2006). To collect this data, I decided to ask the participants of my interview to self-identify their teaching style used in a proof classroom. Active learning has many definitions and encompasses a variety of teaching-styles (Cattaneo, 2017). Further, Cattaneo (2017) found that research studying active learning often has a variety of teaching methodologies used. Rather than constrain my sample to only include instances of project-based learning or other more specific areas within active learning, I instead decided to include the broad category of active learning. To accompany the perceptions the participants of this study hold regarding how their students use examples, this study additionally assumes participants perceive their own teaching appropriately when classified as lecture or active learning. Six participants indicated they used a lecture style of teaching and five participants indicated they used an active-learning style of teaching.

3.3 Data collection

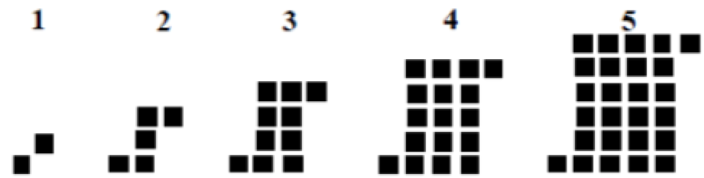
One-hour semi-structured interviews were conducted with all participants online through Zoom. Participants were emailed two questions to be used in the interview in advance of the interview. They were not expected to solve the problems, but this was intended so that less time in the interview would be spent on solving the problem given that they would have previewed the mathematical context. Participants were first asked to provide information used to help determine the characteristics described above. Then, participants were shown one of the two problems for the interview. The two problems were titled the *growing S pattern task*, and the *polygon diagonal problem*. Figures 1 and 2 show these two problems.

Figure 1

Growing S Pattern Task

Growing S Pattern Task

Question 1. Consider the pattern below. How many square tiles would there be in the eighth step of this pattern?



Question 2. Write an expression for the total number of square tiles in the figure at an arbitrary step (n) of the pattern.

Question 3. Prove that your expression is a valid representation of the number of tiles at step n of the pattern. You may use drawings, words, numbers, and/or symbols for your proof.

Figure 2

Polygon Diagonal Problem

Question 1. Given the number of vertices in a polygon, find a method to calculate the number of diagonals of the polygon. Explain why your answer is correct.

For each interview, it was randomized whether participants would see the growing S pattern task or the polygon diagonal problem first. After spending a couple of minutes reading over the problem and doing any work needed to solve the problem, participants were shown three student work samples attempting to solve either the growing S pattern task or the polygon diagonal problem. These student responses were chosen to exemplify students at varying stages of understanding the role of examples in proof. The student work samples chosen for the growing S pattern task were adapted student work taken from students in an ITP at a large public university in the southeastern United States. Student work was altered slightly to make it fit slightly better into each level of the Balacheff (1987) example-use framework. The student work samples for the polygon diagonal problem were adapted from my interpretation when reading work from Balacheff (1988). I imagine the levels of example-use described by Balacheff to be a continuum ranging from a naive empiricism use of examples on the left side to a generic example-use of examples on the right side. The student responses I adapted are intended to be at the range of responses one could imagine for a student exhibiting each level of the example-use framework described by Balacheff. Table 4 shows each of these student responses and their accompanying Balacheff level. Figure 3 shows each student's response.

Table 4

Student Work and Balacheff (1987) Level

Balacheff level	Growing s pattern task student	Polygon diagonals student
Naive Empiricism	Nancy	Aaron
Crucial Experiment	Carla	Eric
Generic Example	Gina	Jacob

Figure 3a

Nancy Response - Naive Empiricism, Growing S Pattern Task

Nancy

Question 2. Write an expression for the total number of square tiles in the figure at an arbitrary step (n) of the pattern.

$$S = n^2 + 1 \in R, n > 0$$

S = number of squares, n = iteration

I noticed this pattern through the first several iterations of the task.

Question 3. Prove that your expression is a valid representation of the number of tiles at step n of the pattern. You may use drawings, words, numbers, and/or symbols for your proof.

Iteration	Squares	Correctness
1	$S = n^2 + 1 = 1^2 + 1 = 2$	Valid
2	$S = n^2 + 1 = 2^2 + 1 = 5$	Valid
3	$S = n^2 + 1 = 3^2 + 1 = 10$	Valid
4	$S = n^2 + 1 = 4^2 + 1 = 17$	Valid
5	$S = n^2 + 1 = 5^2 + 1 = 26$	Valid
8	$S = n^2 + 1 = 8^2 + 1 = 65$	Valid
10	$S = n^2 + 1 = 10^2 + 1 = 101$	Valid

The above expression is valid in each case above. This formula will hold for any value of n following this pattern.

Figure 3b

Carla Response - Crucial Experiment, Growing S Pattern Task

Carla

Question 2. Write an expression for the total number of square tiles in the figure at an arbitrary step (n) of the pattern.

$$[(n + 1) * (n - 1)] + 2$$

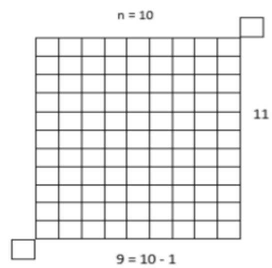
Question 3. Prove that your expression is a valid representation of the number of tiles at step n of the pattern. You may use drawings, words, numbers, and/or symbols for your proof.

n = 1

[(1+1)*(1-1)] +2 = 2

I tried the smallest case and a larger case. Since this formula holds for both, it makes sense for all cases.

n = 10



11 = 10 + 1

9 = 10 - 1

11 * 9 = 99 tiles

99 + 2 = 101 for the total tiles

Figure 3c

Gina Response - Generic Example, Growing S Pattern Task

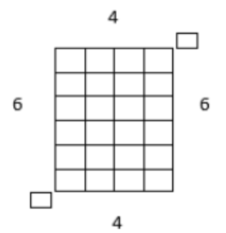
Gina

Question 2. Write an expression for the total number of square tiles in the figure at an arbitrary step (n) of the pattern.

S = number of squares

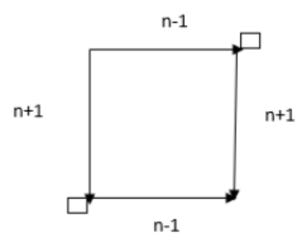
S = $n^2 + 1$

n = 5



For this case, I noticed at step 5 the inside part formed a 4x6 rectangle. The area of the inside is 24 plus the two outside squares for 26 total squares. This is one more than 5 squared.

Question 3. Prove that your expression is a valid representation of the number of tiles at step n of the pattern. You may use drawings, words, numbers, and/or symbols for your proof.



There is an extra block at the top and bottom of each figure. If we take those away, we have a rectangle with an area of $n^2 - 1$ at step n since the side lengths are (n-1)(n+1). When we add the two squares back in, we have $n^2 + 1$ squares at step n.

Figure 3d

Aaron Response - Naive empiricism, Polygon Diagonal Task

Aaron

Question 1. Given the number of vertices in a polygon, find a method to calculate the number of diagonals of the polygon. Explain why your answer is correct.

Well, a polygon with 6 vertices has 6 x 3 = 18 diagonals. In a polygon with 8 vertices, there are 8 x 5 = 40 diagonals.

So, the number of diagonals in a polygon equals the number of diagonals from one vertex multiplied by however many vertices there are.

Figure 3e

Eric Response - Crucial Experiment, Polygon Diagonal Task

Eric

Question 1. Given the number of vertices in a polygon, find a method to calculate the number of diagonals of the polygon. Explain why your answer is correct.

Number of Vertices	**Number of Diagonals**
6	9
7	14
10	?

First, I found that a polygon with six vertices had nine diagonals. Then, a polygon with seven vertices had five more diagonals than that one for a total of fourteen diagonals. This led me to a pattern.

Then, I tried to test the pattern by jumping ahead to a big shape with ten vertices. If my pattern holds, this one should have 35 diagonals. When I tested this by drawing it out, the pattern held. This means the claim is true.

Figure 3f

Jacob Response - Generic Example, Polygon Diagonal Task

Jacob

Question 1. Given the number of vertices in a polygon, find a method to calculate the number of diagonals of the polygon. Explain why your answer is correct.

I started from a polygon with 7 vertices. The number of diagonals to each vertex is n – 3.

For a polygon with 7 vertices, each point has two neighbors. So:

7 – (2 + 1)   The 7 is the number of vertices, the 2 is the number of neighbors, and the 1 is the point itself

In general, n – (2 neighbors + itself)

Then take this and multiply by the number of vertices. This counts each diagonal twice, so divide by 2 for the answer.

So, my final answer is n(n-3) all divided by 2.

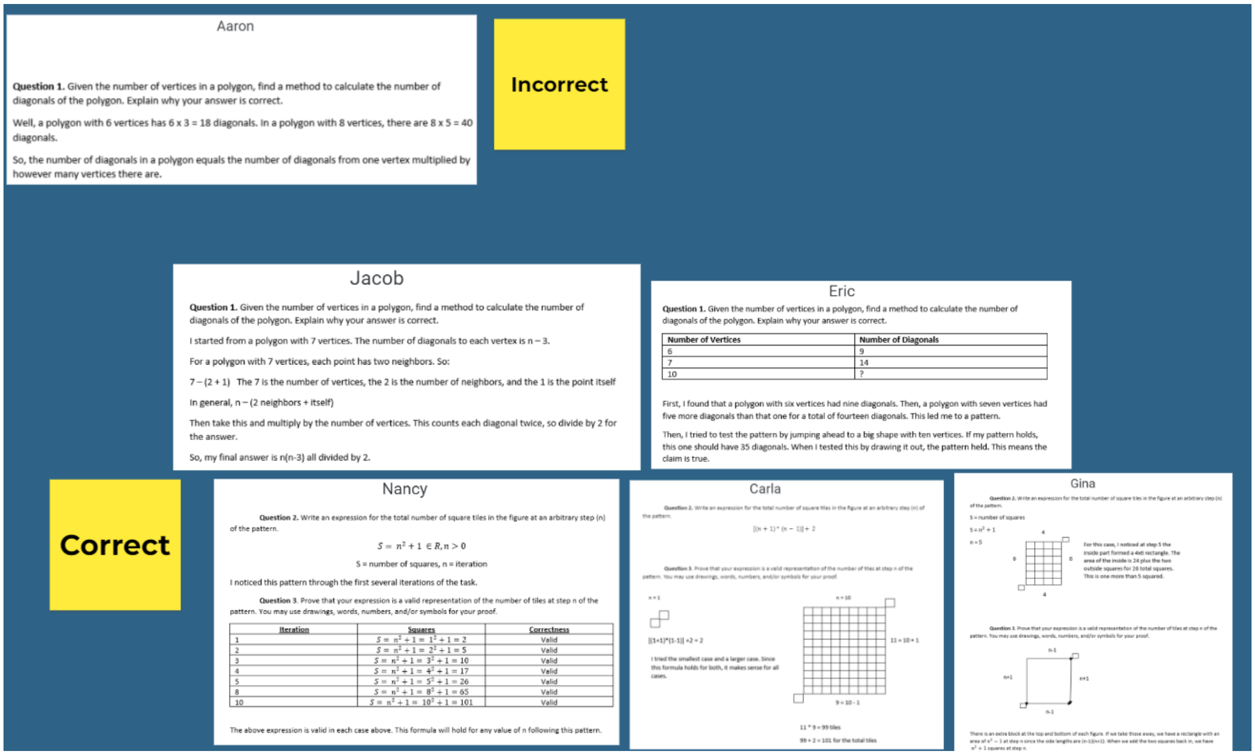
For each interview, the order of each student work sample was also randomized within the growing S pattern task or the polygon diagonal problem, and no labels or categorizations were provided to the participants for the student work except for the students' name (pseudonym). The three student work samples for each problem were always shown together with either three growing S pattern task problems presented in some random order together first or the three polygon diagonal student work samples presented in some random order together first.

Participants were asked first to describe any noticing from the provided student work sample. Participants were informed this could include anything from commenting about what they liked or disliked about the student work provided, how common a proof like this was in their classroom, or more specific feedback about the proof. Participants were then asked to give any feedback or questions they might ask of this student if they were a student in their classroom presenting this student work sample as a homework assignment to be turned in. After each of the questions, I asked participants if they would use a question like the one provided in their ITP class. Although only four participants said they would use or have used the exact questions presented in the interview, ten of the eleven participants indicated they thought the questions were appropriate for an ITP class as substitutes for similar questions they asked.

After participants responded to all six student response samples across the two tasks, they were asked to sort the student response in any way they deemed appropriate. Participants used the share screen feature on Zoom to move around screenshots of the student work samples on a Google Jamboard. There were no constraints to the size of the categories made, number of categories to be made, or criteria for the categories. Participants were provided as many slides on Google Jamboard as they needed to provide as many groupings as they could imagine. Figure 4 details one instance of how a participant sorted the student work in multiple ways.

Figure 4a

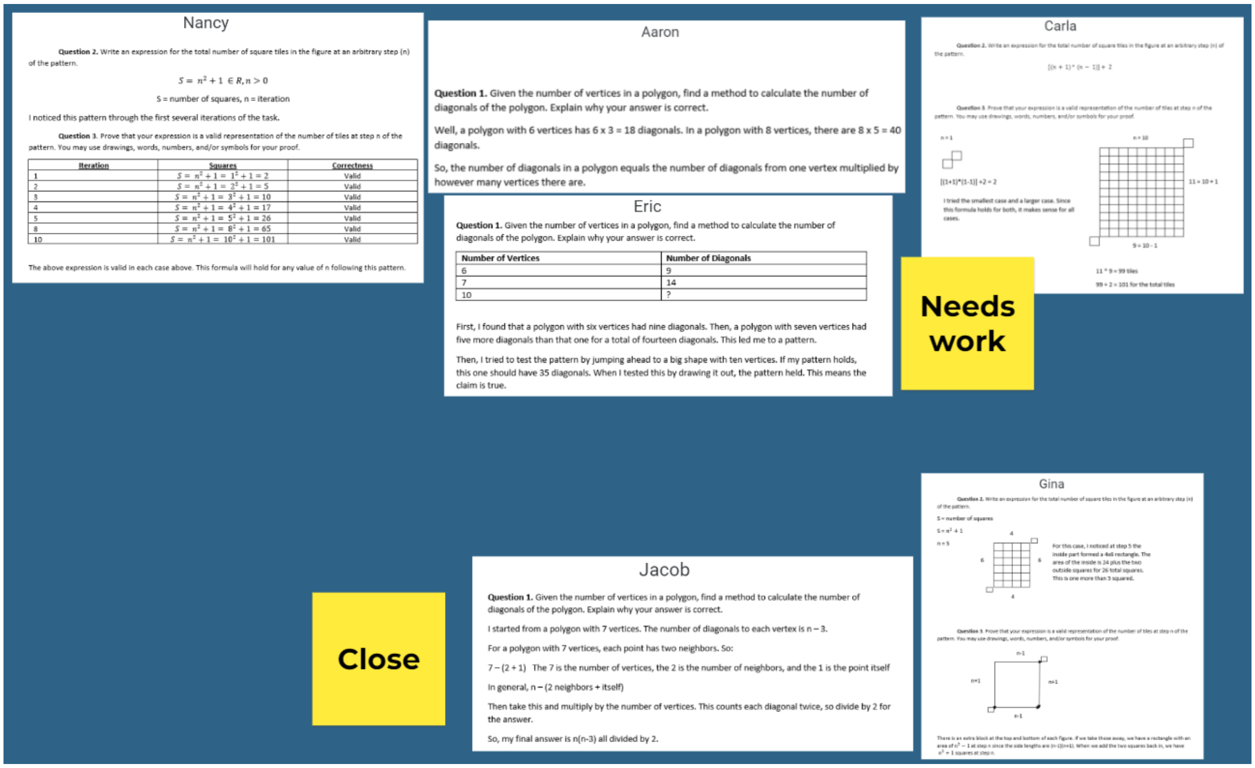
First Sorting Done by Dr. Nathan



Here, Dr. Nathan grouped five student responses together leaving Aaron (naive empiricism, polygon diagonal) alone. Dr. Nathan titled his groups “Correct” and “Incorrect”. In the following picture, Dr. Nathan instead grouped responses by what he perceived the student response to indicate their understanding of the question. The group containing Jacob (generic example, polygon diagonal) and Gina (generic example, growing S) was titled “Close” while the group containing the other four participants was titled “Needs work”.

Figure 4b

Second Sorting Done by Dr. Nathan



Next, participants were asked to rank the six student responses from one to six. A one signified the best attempt at a proof, while a six signified the worst attempt at a proof. Participants were asked to provide reasoning for each students' placement. Finally, participants were asked for specific feedback to be given to the student they ranked a one and the student they ranked a six if the student were in their classroom and provided the student work sample they were shown.

3.4 Analysis

Data in this study comprise transcriptions of the interviews as well as screen captures of the sorting and ranking tasks. For the ranking task, I made a table of how each participant ranked the six student responses and looked for trends among characteristics of the participants. For the sorting task, I looked first for identical sorting done by multiple participants. These identical sortings were categorized in a group together. Next, I looked at the names participants gave their sortings and included some groupings that were not identical but expressed similar reasoning with a similar title. Finally, I looked for any other commonalities amongst sortings such as students commonly grouped together, or students commonly left alone in their own group.

To look for themes amongst participants' responses in the interview, I looked at the transcriptions of the videos. The recorded Zoom interviews were transcribed by talk turn. I define a talk turn as all the words spoken by one participant of the interview until interrupted by another. These interruptions were typically when the interviewer asked the participant to move to another question or asked clarifying questions of something the participant said. Using talk turns in this way allows me to include the context I deemed relevant to the story of the interview. After transcribing the interviews, I first looked for themes within each participant's responses.

Themes could be anything ranging from common instructional responses to student work to common noticings of how students think about a problem. I checked for themes in an open coding style without any prescribed notion of what I might find. If a participant responded with a similar response or noticings across at least two different student work samples, I considered this a potential theme. For instance, Dr. Amanda's transcript included a potential theme of inductive thinking. Dr. Amanda remarked when responding to Carla, "Okay, so here's some induction ideas. You've got the base case. And a small case for largest." Similarly, when Amanda was responding to Nancy's work, "[Nancy] tried a smallest case. And the larger case, is suggesting that there's maybe some idea there of like, maybe I can interim check intermediate cases, which sort of feels like you're building to induction." Dr. Amanda further noted similar thoughts about Gina and Eric's work. This was sufficient to be a potential theme within Dr. Amanda's transcript.

I took sections of the transcription that corresponded to each potential theme from all participants and checked these against each other. To be considered a theme across the participants, I only considered potential themes that were shared by at least three participants. Since only two participants had potential themes related to inductive reasoning, I did not count looking for inductive reasoning as a theme in this work. One instance of a theme across participants is the theme of inquiry about starting position. This theme arose when participants would mention wonderings about why a student decided to start with a particular case or number in mind. For instance, Dr. Nathan remarked about Jacob, "The only thing I'm curious about is why. Why did you start from a polygon with seven vertices? Why not five or four?" Similarly, Dr. Nathan questioned Eric, "Why six? Why seven and 10? I don't know why these numbers were chosen." Dr. Nathan had similar responses for both Nancy and Carla, leading me to consider inquiry about starting position a potential theme. I decided to consider inquiry about starting position a theme across participants when Drs. Amanda, Sarah, and Phil had similar potential themes to Dr. Nathan about why the students chose to start at specific numbers in their proofs.

4. Results

Recall my research question: How do instructors of ITP classes perceive students' understanding and use of examples? I plan to answer this question in four steps. First, I will list how the participants ranked the student responses. This ranking gives validity to the participants intuitive alignment with Balacheff's (1987) example-use framework. Next, I will discuss the common ways my participants grouped the student work in the grouping task. These groupings give some insight into the similarities of how instructors perceive the student work shown to them. Then, I will discuss the degree of acceptance of example-use from the participants. Finally, regardless of the outcome of acceptance of example-use, participants gave similar responses for how to move the students forward. This will comprise the fourth section of my findings.

4.1 Rankings

I found it promising that all 11 participants, when asked to rank the student responses, ranked the two generic example students (Jacob and Gina) in the top three. Further, 10 of the 11 participants ranked these two students either first or second. The one participant who ranked Jacob (generic example - polygon diagonals) third remarked that due to the difference in nature between the questions he might be placing more emphasis on what Carla (crucial experiment - growing S) did (whom he ranked second) than Jacob. Similarly, 10 of the 11 participants ranked the two naive empiricism students (Aaron and Nancy) in the bottom three. The participant who ranked Aaron (naive empiricism - polygon diagonals) third remarked that he thought Carla may have used outside sources such as a friend or textbook to aid her proof and chose to rank her lower than other participants. Table 5 shows the rankings of each student across all 11 interviews.

Table 5

Rankings of Student Work by Instructors

	1st	2nd	3rd	4th	5th	6th
Gina (generic example)	10	1				
Jacob (generic example)	1	9	1			
Carla (crucial experiment)		1	9	1		
Eric (crucial experiment)					8	3
Aaron (naïve empiricism)			1	5		5
Nancy (naïve empiricism)				5	3	3

These rankings show that practitioners of ITP courses align with the levels described by Balacheff (1987) example-use framework. When adapting the student work samples, I was primarily focused on the mathematical concepts involved with the student reasoning. Although Eric (crucial experiment - polygon diagonal) was positioned to be a student exhibiting the crucial experiment level of example-use as described by Balacheff , most participants held issue with the statements, "This led me to a pattern." and, "If my pattern holds." Participants routinely expressed concern with a potential gap in knowledge shown by Eric excluding his pattern from the written work. For instance, Dr. Nathan commented, "I'm unclear what pattern you (Eric) are talking about here. (Reading Eric's work to himself) When I tested it by drawing it out, the pattern held. What? How did you test it? What pattern?" Comments like these led Eric to be consistently ranked lower than at least one of the students who were adapted to show a naive empiricism level of understanding.

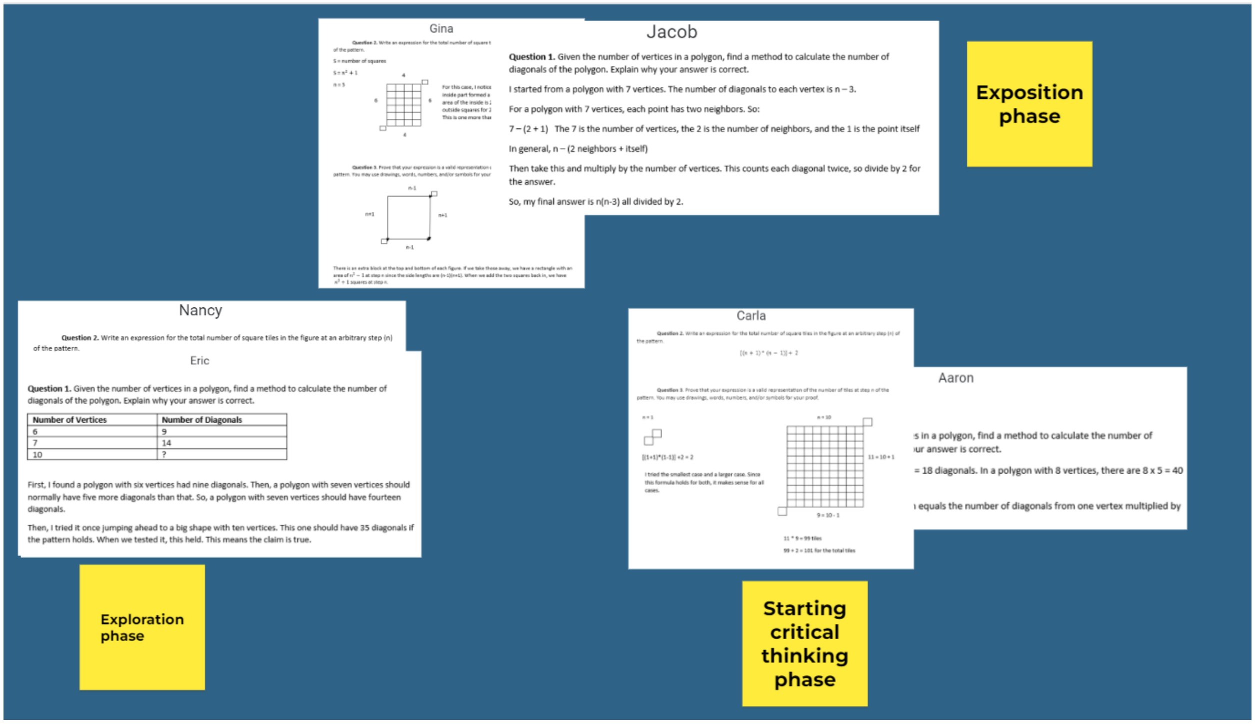
4.2 Common groupings of student work

After analyzing the grouping task participants completed in the middle of the interview, two common categories emerged. Groupings are reported only when done by at least three participants explained in a similar manner. Five participants created a single grouping for the student work. Three participants created two unique groupings, and three participants created three unique groupings. This provided a total of 20 groupings. Several participants who grouped in multiple ways are included in both common categories.

The first and most common grouping was the perceived understanding grouping. This grouping was done by 10 of the 11 participants. Most of the participants who used this grouping used terminology such as "needs improvement" or "still figuring things out" to denote what they perceived to be the lowest level of understanding. Participants used "approaching understanding" and "critical thinking" to denote what they perceived to be the middle level of understanding. Finally, participants used terminology including "near completion", "A students", and "high level thinking" for what they perceived to be the highest level of understanding. Figure 5 shows an instance of this grouping.

Figure 5

Grouping by Perceived Level from Dr. Phil



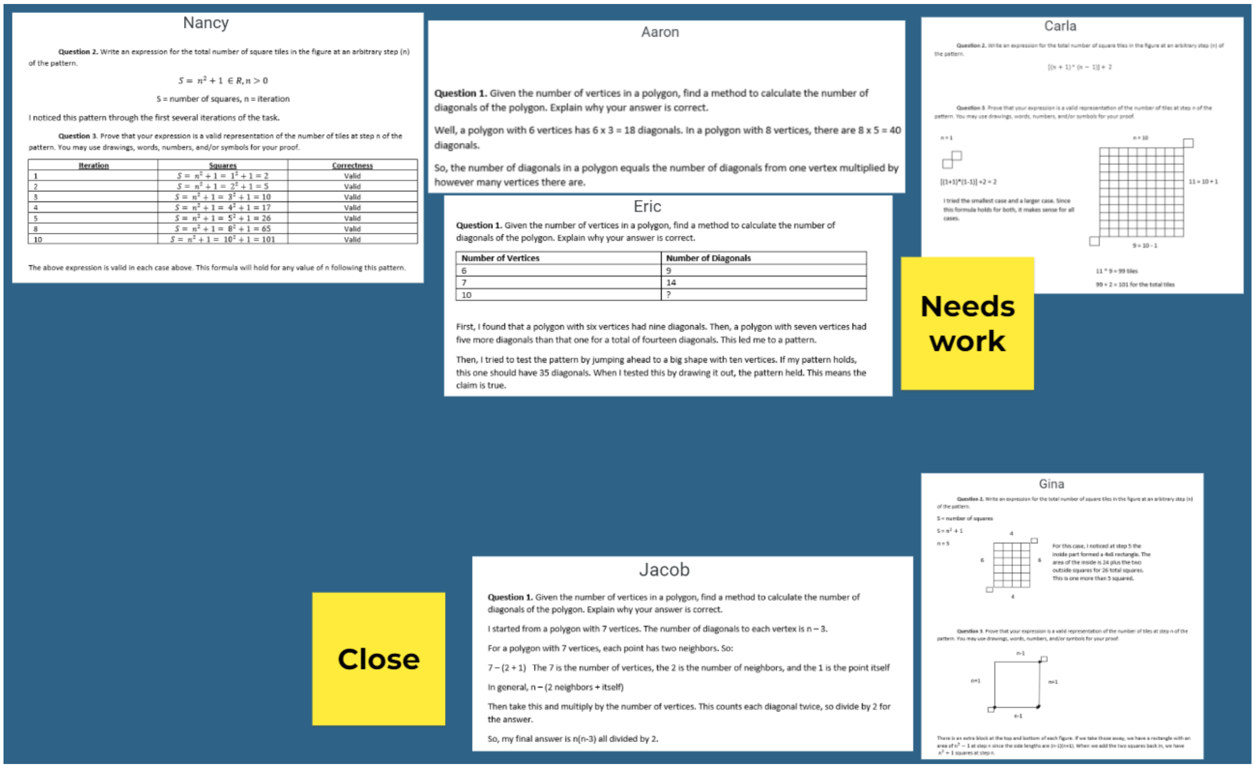
The factors that determined this sorting were the inclusion of three separate categories and explicit discussion related to the understanding shown by the student in the problem. Four of the instances of this grouping looked identical to the picture above with slight variations in the names given by participants for the groups. In all ten instances of this grouping done by participants, Gina and Jacob were in the highest grouping. Three of the instances of this grouping also included Carla in the highest level of understanding. In all ten instances of this grouping done by participants, Eric was in the lowest grouping. As mentioned above, the unclear wording in Eric's response is likely the reason most of the participants put Eric in a lower category. This was evidenced by Dr. Chris commenting on why he grouped Eric in the lowest perceived grouping, "Eric did some things, but I wasn't sure what the heck he was doing." Similarly, Dr. Phil gave the following excerpt about grouping Eric and Nancy together,

> I'm going to group Eric and Nancy. Because I think they're on the same track here. They're both in what I think was a very exploratory phase of mathematics. In Nancy's case, just sort of looking at numbers and see if she can see something that makes sense to her. Eric is just experimenting. I'm looking at chaos.

The second common grouping of student work samples was the approval category. This grouping was done by six of the 11 participants. The requirements for this category included explicit discussion about whether the instructor would be happy with this assignment turned in as a homework assignment in their class. Five of the six participants who did this grouping included only Jacob and Gina in the "accept" category. The sixth participant also included Carla in the "accept" category. Two participants grouped the other four students all together in the "reject" category. The other four participants had two or three additional categories for levels of rejection. In each of these cases, Carla was in the category closest to the top while Eric was in the category at the bottom. Figure 6 details an instance of this category.

Figure 6

Acceptance Categorization by Dr. Nathan



These groupings show many participants intuitively group student work in alignment with the levels of example-use described by Balacheff (1987). Ten of the 11 participants grouped the students at the generic example level of example-use together in a category. The one participant who didn't group student work in the grouping titled perceived understanding put Gina in her own category as the student who was the most advanced. This participant then grouped based on the approach to the problem with students who used a table (Eric and Nancy) together and the other students separately.

4.3 Acceptance of example-use in proof

Of the 11 interviews conducted, five participants expressed concern with seeing examples of any form written in the work shown by students. The other six participants expressed explicit recognition and excitement to see the students using examples to aid their reasoning. Table 6 describes the demographic information of the participants who disliked examples in the student work. Table 7 describes the demographic information of the participants who liked the students using examples.

Table 6

Demographic Information of Participants who Disliked Examples

Participant	Dr. Daniel	Dr. Amanda	Dr. Sarah	Dr. Hubert	Dr. Casey
Carnegie Classification	M1	Baccalaureate College (Diverse Fields)	R1	R1	Baccalaureate College (Arts & Sciences Focus)
PhD received	Mathematics	Mathematics*	Mathematics Education	Mathematics	Mathematics
Length of time at university	9 Years	6 Years	5 Years	30 Years	19 Years
Private vs. Public university	Public	Private	Public	Public	Private
Self-described Teaching style	Active learning	Active Learning	Active Learning	Lecture	Lecture

* This participant expressed they have published in mathematics education research. Note all Carnegie classifications come directly from the website carnegieclassifications.acenet.edu

Table 7

Demographic Information of Participants who Liked Examples

Participant	Dr. Phil	Dr. Nathan	Dr. Tucker	Dr. Chris	Dr. Ryan	Dr. Joe
Carnegie Classification	R2	Baccalaureate College (Diverse Fields)	M1	M1	D/PU	Baccalaureate College (Diverse Fields)
PhD received	Math*	Math Education	Math	Math Education	Math Education	Math Education
Length of time at university	18 Years	4 Years	7 Years	18 Years	20 Years	1 year
Private vs. Public university	Public	Private	Public	Public	Private	Public
Self-described Teaching style	Lecture	Lecture	Active Learning	Lecture	Lecture	Active Learning

* This participant expressed they have published in mathematics education research.

Participants were classified as either broadly speaking "liking" examples or "disliking" examples based exclusively on their unprovoked remarks during the first phase of the interview responding to student work. All participants repeated their preference (either liking or disliking) for examples at least twice throughout the interview. Descriptions are given of participants in each classification of example-use in the following section. By "liking" and "disliking", I mean explicit comments made during the interview of whether an example (productive or unproductive) should be included in the student work samples. Participants were informed that the student work samples were meant to be turned in as homework assignments early in the semester of an

ITP course. Participants in the "liking" examples category made clear they appreciated seeing the student reasoning through examples in multiple of the student work samples. Further, participants in this group expressed a desire to see more reasoning from students that were ranked lower in Table 5. Participants in the "disliked" examples category expressed concern for seeing the examples in the written work to be turned in as a homework assignment.

Dr. Tucker was classified as liking examples by responding to the student work from Jacob (generic example - polygon diagonal), "So, I do like the, I guess, [he] kind of started out with, you know, thinking of a fixed number. Just to kind of get an idea of what to do with it (the problem)." Similarly, Dr. Tucker later gives similar remarks for both Gina (generic example - growing S) and Carla (crucial experiment - growing S). This type of remark about examples for both Gina and Jacob was shared by every participant in the liking examples category. Five of the six participants in this group additionally raised concern when examples were used by Nancy (naive empiricism - growing S) and Aaron (naive empiricism - polygon diagonal) in what could be classified as an unproductive manner (Aricha-Metzer & Zaslavsky, 2019). For instance, Dr. Tucker said when referring to Eric and Aaron,

> I guess generally, a pitfall I see with students is trying to just do a couple of examples. And then that kind of just, I know this doesn't say prove (referring to the wording of the question given), but kind of thinking that counts as a proof. And yeah, it feels like they just kind of need to bring more justification of why would this work in general.

This provides evidence that these instructors implicitly approved of examples when used productively but were cautious of examples when used unproductively.

Participants in the "disliked examples" category responded to student work of all levels with similar wording like Dr. Hubert referring to Gina, "I mean, it's reasoning by example. Which is always rough. I would say that the n minus one, n plus one part could have helped me earlier on." Dr. Hubert continued to explain he thought Gina's argument was weaker with the inclusion of the answer to question 2 showing how she came up with her formula. Rather than include a section of her reasoning, Dr. Hubert expressed an expectation of formality in the written product from Gina, "So I'm struggling right now with if you can do this (referring to Gina's answer at the bottom of the page), how does it not stick? I mean, I ought to be able to state that part at the end earlier, right?" Similarly, Dr. Sarah raised concerns with how Jacob approached his proof,

> Okay, so it's interesting, because like, I think Jacob is thinking about a seven-sided figure, right? And he is taking it as like a generic example. But then the seven sided figure is actually never important in his work.

Unlike the liking examples category, participants in this category expressed concern with all six students' work with regards to the use of examples.

When looking between the groups of participants who expressed concern for all examples and those who liked generic examples, there are almost no differences. There is an even spread of participants in both groups from every Carnegie classification, public versus private college, self-described teaching style, and varying lengths of time teaching at the university. There is a potential difference between the two groups with a tendency for the participants with a PhD in mathematics education to prefer generic examples. Four of the five participants with a PhD in mathematics education were classified in the liking examples category. Of note, two participants in my study with a PhD in mathematics expressed an interest and active participation in mathematics education research. These two participants were split in their preference of liking examples.

Following discussion of whether the participants liked or disliked examples, most participants included some rationale behind their reasoning. Largely, the reasoning provided by participants related to their conceptions of the role of the course they were teaching. For instance, Dr. Amanda commented, “There's really good ideas in all of them (referring to student work as a whole), but they need a little bit of massaging before they're, you know, ready for publication.” Dr. Amanda continued to suggest that one of her goals for the course was to prepare students to write proofs that could be “publishable.” Participants were not explicitly asked about their conceptions of an ITP course or of proof in general, however ideas of formality and rigidity were shared amongst all participants in the disliked examples group. Dr. Chris remarked in his final thoughts of the interview,

> With beginners, I want you to be overly verbose and showing why you know you're right. If you're a PhD student, it's funny, I would let a PhD student in math get away with writing a proof where you leave steps out, or you gloss over certain details, because at that point, you're an expert in the language, I'm not so much worried about every detail. With learners, I worry more about whether you have every step down.

Dr. Chris continued to use this line of reasoning as a foundation for why he believed seeing the examples and reasoning written down in his students' work was essential to the learning process.

4.4 Instructional feedback

The final part of the interview had participants give instructional moves or advice to respond to the students who they ranked first and sixth in the ranking task detailed in

Table 5. Ten of the 11 participants ranked Gina (generic example - growing S) first. The other participant ranked Jacob (generic example - polygon diagonals) first. Some participants gave instructional advice for multiple candidates outside of their first and sixth ranked student. All five participants in the "disliked examples" category had almost identical responses to either Gina or Jacob. Dr. Amanda remarked about Jacob (whom she ranked second):

> So, I think that I would push him to sort of argue this without saying, I started from seven and blah, blah, blah, to make sure that it really is cemented in. Okay, but we can extend this indefinitely. So, I would probably ask him something simple, like, well, okay, but if you had started with eight vertices, what would have happened? And can you then extend that without relying on the number of vertices that you start with?

Similarly, Dr. Sarah remarked about Jacob (whom she ranked second), "The fact that Jacob does still have, like the thinking about his example. To me, I would not want to see evidence of it in the formal proof."

Conversely, all six instructors in the "liked examples" category gave similar ideas to either Jacob or Gina to consider if they could prove the given task in an alternative method. Dr. Phil remarked about Gina,

> I would want to see Gina, can you come up with a different kind of argument? If I thought it would be productive I might even leave that open ended. Because that's (referring to Gina's picture) one way to decompose the picture and Gina has had success. But there are other ways to decompose the picture.

Nine of the 11 participants in the interview had similar recommendations to the students they ranked at the bottom of their list spanning both the liked examples and disliked examples groups. These participants commonly expressed a desire to see these students draw a picture and connect the picture to a general argument to build intuition. Dr. Hubert remarked about Eric (whom he ranked sixth), "I would certainly say to back this argument up a bit more, draw the picture to show me where the diagonals are."

Similarly, Dr. Tucker asked of Nancy (whom he ranked sixth),

> Just kind of push her towards looking back at this picture, and trying to connect together, you know, where these numbers are coming from. And so the problem is the n squared thing is harder to see if you don't already have it figured out in terms of like an n minus one, n plus one.

The final two participants who had a different recommendation to students ranked at the bottom were split across both the liked examples and disliked examples group. Both participants recommended changing the numbers of a problem to larger numbers to discourage the tendency of proof by example. Dr. Amanda commented on Nancy,

> But what if the number we needed to check was like, a million or something that is like so obnoxiously large that they couldn't? You know, you could do the number but maybe you couldn't, You couldn't attack it with a picture. So then how do you sort of pushing them to think abstractly because it's too large to sit down and try to calculate.

Most participants encouraged students they viewed to be struggling with the proof to use examples to aid their understanding. Some of the wording from these participants like the quote from Dr. Tucker indicate how to connect a specific example to the reasoning behind the problem akin to a generic example argument. These recommendations for struggling students were seen across both the liking examples and disliking examples groups.

5. Conclusion and discussion

In this study, I sought to lay groundwork to answer a call to research from Zaslavsky and Knuth (2019). The field of example-use in proof has a great foundation for understanding how and why students use examples to aid their proof production and understanding. However, there is much less research on how to aid instructors of ITP classes to help their students use examples more effectively. I aimed to take a first step to fill in this gap by first investigating what instructors of ITP courses think when faced with students at various levels of example-use. My findings can assist future researchers in developing instructional interventions for practitioners in ITP courses by considering their acceptance of example-use. Considering the acceptance or rejection of example-use will be beneficial when designing professional developments. Desimone and Garet (2015) described how the most productive professional development involves understanding the beliefs and knowledge of the participants to meet them at an appropriate level. This study helps indicate the current awareness of practitioners' knowledge related to example-use as well as their beliefs surrounding the effectiveness of example-use.

My research question was: How do instructors of ITP classes perceive students' understanding and use of examples? I answered this question in two ways. First, instructors largely align their perceptions of the usefulness of examples with the framework described by Balacheff (1987). This was shown through both the rankings of student work and groupings done by instructors who often grouped student responses based on how the instructors perceived their students to understand the material. Second, there is a split amongst practitioners of ITP courses on whether examples should be included in the written proofs produced for coursework. Some participants in my study appreciated the clear showing of student thinking in the homework samples, while other participants in my study preferred to see more formal writing that excluded examples. Regardless of this split, however, participants often suggested similar instructional activities to students they perceived to be struggling, including the explicit use of examples to aid their intuition in approaching a proof.

As a field of mathematics education researchers, we have argued productive examples are beneficial for the learning and production of proofs (e.g., Aricha-Metzer & Zaslavsky, 2019; Balacheff, 1987; Ellis et al., 2019). This interview task was designed to include two student work samples (Jacob and Gina) who used examples at the generic example understanding as described by Balacheff (1987). When students used examples generically and gained insight into how to approach the proof, Aricha-Metzer and Zaslavsky (2019) argued these were productive instances of examples. Aricha-Metzer and Zaslavsky further cite Leron and Zaslavsky (2013) and argued using generic examples, "may be a worthwhile pedagogical tool for teachers in creating engaging activities for students to develop an understanding of proof and proving." (p. 321). I raise concerns that some practitioners of ITP courses might not see the same merit in generic examples.

It is promising that most of the participants ranked students aligning with the levels of example-use described by Balacheff (1987). This provides insight into the current knowledge held by practitioners of ITP courses to help future professional development be as effective as possible. One discrepancy that came up in this ranking dealt with Eric. Leaving remarks in student work such as, "I noticed a pattern," seemed to prioritize much of the discussion surrounding this student. Future work should investigate student work samples where other arguments are made to varying degrees of effectiveness with the student leaving out details as Eric did.

Aricha-Metzer and Zaslavsky (2019) argued generic examples aided students' ability to understand proofs and learn to prove. Further, Aricha-Metzer and Zaslavsky noted almost half of the cases in their study saw students use examples naturally and without instruction from an interviewer. In my study, many participants took issue with the example being in the written product presented to them. Many participants in this study intuitively confirmed the difference between productive and unproductive example-use described by Aricha-Metzer and Zaslavsky. Ten of the 11 participants in this study grouped the two students (Jacob and Gina) who were at the generic example level of understanding as described by Balacheff (1987). Several participants even grouped these two students together as an "accepted" grouping of an argument they thought was beneficial to the student.

Participants were split on acceptance of examples of any kind in student work. Five of the participants expressed how they wished students who were using examples productively would take this reasoning out and instead put together what they deemed an appropriate response. Dr. Sarah's comments about Jacob's illustrate one side of the split, while comments like Dr. Tucker's highlight the other. Recall Dr. Sarah noted, "The fact that Jacob does still have, like the thinking about his example. To me, I would not want to see evidence of it in the formal proof." I encourage future research to investigate what personal views practitioners hold regarding the purpose of an ITP course.

I did not find evidence to support a claim that teaching experience impacted the acceptance or rejection of example-use. Similarly, I found no evidence regarding Carnegie classification of institutions, public versus private institutions, or self-described teaching style. There was a slight leaning of participants who had a Ph.D. in mathematics education to express their thoughts leading to being classified in the liking examples category.

Despite the differences between acceptance of written examples in a product to be turned in, many participants shared a common instructional practice to students they deemed struggling. Nine of the 11 participants in my study expressed they would like to have the students they ranked as struggling draw a picture or use a specific example to help gain intuition for how to approach the problem. A common strategy used by instructors of ITP courses from this study is to encourage their students to use examples to help aid their understanding of proof when the students are stuck. After the students use these examples to help generate these ideas, some instructors then expect that students will omit the informal example-based reasoning from their formal proof submissions. Mathematics education researchers should consider whether students see and understand this distinction that is made by some professors. It is important in our messaging to students to be consistent with respect to the utility and presence of examples in proving.

This study is not without limitations. First, the two problems chosen (growing S pattern task and polygon diagonal task) were not familiar to many of the participants. These tasks were chosen to be easily accessible to the instructors so most of the interview time would be spent on the student solution rather than solving a more difficult proof. Although most participants expressed that they believed these tasks were appropriate for students in an ITP course, the nature of seeing a problem for the first time may have affected the responses of several of the participants.

Secondly, I did not show all types of proof to the participants. It could be the case that all participants strongly support the use of generic examples in proof methods like induction but dislike the use of examples in other methods like the ones included in my study. Third, all six student work samples used examples to some degree in this interview. Some participants may have expected to eventually see a more formal "correct" proof at some point in the interview and expressed disapproval for the proofs presented in expectation of another method to come later.

I do not claim these results to be generalizable to instructors of ITP courses. I do, however, wish to draw attention to the instructor's perceptions of the purpose of both ITP courses as well as the nature of proof itself. Future work would benefit from investigating these perceptions.

Many future directions of this research come to mind. First, future research should investigate how the acceptance or rejection of examples are tied to the proof method employed by the student. Next, research is needed on the relationship between instructors' conceptions about the nature of mathematics and formality involved in mathematical proof and their acceptance or rejection of example-use. I have preliminary findings suggesting the participants' expectations about an ITP course relate to their acceptance of example-use. However, further research is needed to determine how participants reach their conclusions. Finally, the goal of developing instructional interventions as described by Zaslavsky and Knuth (2019) should take into consideration the acceptance or rejection of example-use by practitioners of ITP courses.

Article 2: Instructors Responses to Student Work Using Examples (Intended for Journal of Mathematical Behavior)

1. Introduction

Many mathematics education researchers have investigated the struggles students often have with proving (e.g., Aricha-Metzer & Zaslavsky, 2019; Stylianides & Stylianides, 2009; Weber, 2010). A common finding is that students often misunderstand the role of examples in a proof context (e.g., Balacheff, 1987; Harel & Sowder, 1998, 2007; Healy & Hoyles, 2000). In an introduction to proof (ITP) course this is particularly relevant when students are learning about proof for possibly the first time (Boyle et al., 2015). This has led to a push from researchers in recent years to investigate the role of example-use by students in proof classes (Ellis et al., 2019; Zaslavsky & Knuth, 2019). Although much has been understood about the manner through which students use examples (i.e., Ellis et al., 2019; Aricha-Metzer & Zaslavsky, 2019), far less is understood about how to help instructors get their students to use examples more effectively. Zaslavsky and Knuth (2019) made a call for action to look at how to provide instructors of proof classes materials and support to help their students use examples more productively.

Before designing an instructional intervention or instructional materials to help instructors of proof classes guide their students in the use of examples, it is important to understand what instructors currently do in their own practice. This study investigates the current practices of instructors of ITP across the Southeastern United States. By understanding what instructors of ITP courses are currently doing and how they respond to students' use of examples in a proof setting, future work can design instructional interventions or tasks to help these instructors guide their students to use examples in a more effective manner.

1.2 Research question

The purpose of this study is to understand how practitioners of ITP courses respond to samples of student work. These student work samples highlighted students using examples to aid their proof production to varying degrees of success. These participant responses were prompted in an interview with, "If these students were students in a class you were teaching and submitted this work as a homework assignment, how would you help the student move forward?" I seek to gain understanding of how practitioners are currently responding to student work to help better prepare future instructional interventions aimed at improving practitioners' skills at helping their students use examples more productively. To help with this understanding, I pose the following research question: In what ways do instructors respond to student work samples involving examples in an ITP course?

2. Theoretical Framing

This study is grounded in theoretical framing in two parts: (1) example-use in proof and, (2) mathematical knowledge for teaching proof. Example-use in proof helps position how students can come to understand the purpose and use of examples in a proof setting. Further, I will present a framework that guided the development of the interview protocol used to collect data for this study. Second, mathematical knowledge for teaching proof (MKT-P) helps mathematics education researchers understand the facets involved in both teaching and learning proof. Describing the parts of MKT-P will allow me to give further clarity for why I chose to analyze my data using this framework as a guide.

2.1 Example-use in proof

To frame the task used in my interviews for this study, I will describe the Balacheff (1987) framework on example-use. Balacheff described three categories of how students use examples to help guide their mathematical thinking: naive empiricism, crucial experiment, and generic example. A student exhibits a naive empiricism understanding of the use of examples when they try to make a deductive argument or claim based on a small number of cases. A crucial experiment level of understanding of example use goes one step beyond this and instead is built on some reasoning about the specific cases chosen to exemplify multiple aspects of a problem. Finally, the generic example level of example use can be seen in students' reasoning when the reasoning gained from an example leads to a general argument abstracted from the specifics of the argument made in the example. I interpret this framework to be on a spectrum of how students use examples ranging from naive empiricism on the left and generic examples on the right, as depicted in Figure 7.

Figure 7

Continuum of Example-Use

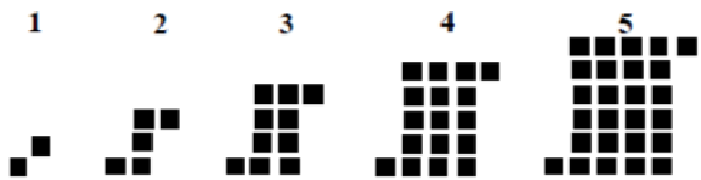
To give further clarity to my understanding of each of these levels, consider the growing S pattern task, detailed in Figure 8, which asks students to identify a generalized expression for the given pattern and to prove their expression is valid. Figure 9a, 9b, and 9c depict varying student work samples for the growing S pattern task. These student work samples encompass a wide range of uses of examples to aid the students' reasoning in proof ranging from naive empiricism to generic example.

Figure 8

Growing S Pattern Task

Growing S Pattern Task

Question 1. Consider the pattern below. How many square tiles would there be in the eighth step of this pattern?



Question 2. Write an expression for the total number of square tiles in the figure at an arbitrary step (n) of the pattern.

Question 3. Prove that your expression is a valid representation of the number of tiles at step n of the pattern. You may use drawings, words, numbers, and/or symbols for your proof.

Figure 9a

Nancy - Naive Empiricism, Growing S Pattern Task

Nancy

Question 2. Write an expression for the total number of square tiles in the figure at an arbitrary step (n) of the pattern.

$$S = n^2 + 1 \in R, n > 0$$

S = number of squares, n = iteration

I noticed this pattern through the first several iterations of the task.

Question 3. Prove that your expression is a valid representation of the number of tiles at step n of the pattern. You may use drawings, words, numbers, and/or symbols for your proof.

Iteration	Squares	Correctness
1	$S = n^2 + 1 = 1^2 + 1 = 2$	Valid
2	$S = n^2 + 1 = 2^2 + 1 = 5$	Valid
3	$S = n^2 + 1 = 3^2 + 1 = 10$	Valid
4	$S = n^2 + 1 = 4^2 + 1 = 17$	Valid
5	$S = n^2 + 1 = 5^2 + 1 = 26$	Valid
8	$S = n^2 + 1 = 8^2 + 1 = 65$	Valid
10	$S = n^2 + 1 = 10^2 + 1 = 101$	Valid

The above expression is valid in each case above. This formula will hold for any value of n following this pattern.

Figure 9b

Carla - Crucial Experiment, Growing S Pattern Task

Carla

Question 2. Write an expression for the total number of square tiles in the figure at an arbitrary step (n) of the pattern.

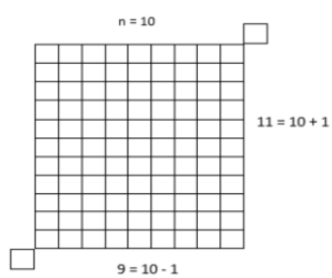
$$[(n + 1) * (n - 1)] + 2$$

Question 3. Prove that your expression is a valid representation of the number of tiles at step n of the pattern. You may use drawings, words, numbers, and/or symbols for your proof.

n = 1

[(1+1)*(1-1)] +2 = 2

I tried the smallest case and a larger case. Since this formula holds for both, it makes sense for all cases.

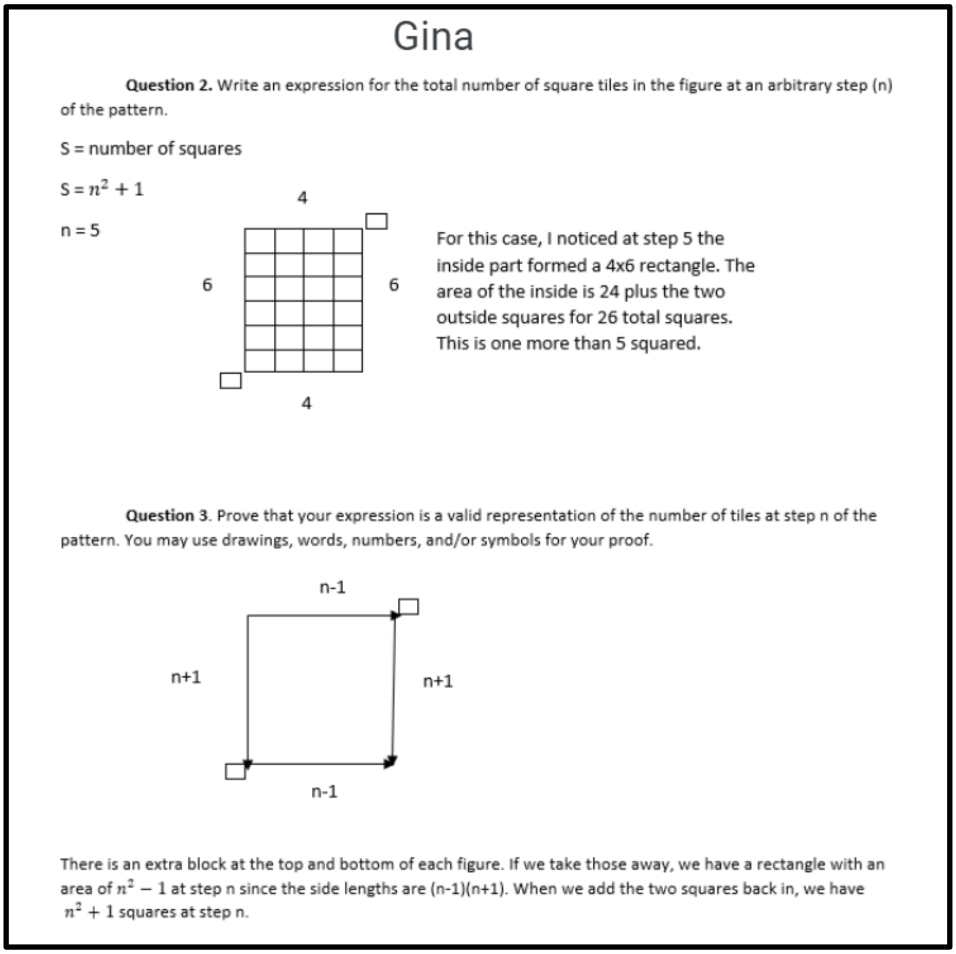


11 * 9 = 99 tiles

99 + 2 = 101 for the total tiles

Figure 9c

Gina - Generic Example, Growing S Pattern Task


Gina

Question 2. Write an expression for the total number of square tiles in the figure at an arbitrary step (n) of the pattern.

S = number of squares

S = $n^2 + 1$

n = 5

For this case, I noticed at step 5 the inside part formed a 4x6 rectangle. The area of the inside is 24 plus the two outside squares for 26 total squares. This is one more than 5 squared.

Question 3. Prove that your expression is a valid representation of the number of tiles at step n of the pattern. You may use drawings, words, numbers, and/or symbols for your proof.

There is an extra block at the top and bottom of each figure. If we take those away, we have a rectangle with an area of $n^2 - 1$ at step n since the side lengths are (n-1)(n+1). When we add the two squares back in, we have $n^2 + 1$ squares at step n.

In Figure 9a, Nancy presents a list of items to check the validity of her given formula. She asserts the formula presented in her solution to question 2 is correct. For the proof of this claim, she simply checks many values. Since all these values work when plugged in, Nancy asserts this formula must be correct. This is an instance of the naive empiricism level of example-use through the argument (or lack thereof) by Nancy in checking a small number of cases to assert a claim is true. Contrasting the approach of Nancy with that of Carla, Carla shows a crucial experiment level of example-use through the careful consideration of specific cases. Carla is concerned with both smaller cases and larger cases working rather than just plugging in the first few counting numbers, like Nancy. This type of reasoning shows an advancement made by the student toward thinking about validation of the problem, however Carla still draws a conclusion based only on a small number of values she plugged in to verify her formula.

At the far end of the continuum, Gina's work is an instance of a generic example. Gina answers the second problem and uses her knowledge of substituting a number to the formula to build intuition. Rather than verify this formula by plugging in other numbers for her solution to question 3, Gina instead uses the intuition she gained from question 2 to build an argument in question 3 that does not rely on the specific number she used for her example.

Regarding my interpretation of a generic example, I do not interpret a generic example to be a noun or specific object. Instead, I interpret a student using an example generically in the sense of making a deductive argument abstracted from the specifics used in the example. In this way, a student may use an example generically while another student may use the same example but not take the extra step of drawing any generality or deductive argument from the example used. Aricha-Metzer and Zaslavsky (2019) argued, "In this sense, a generic example can be seen as a bridge for students, as it may help them move from empirical views and misleading intuitions to a justifiable conviction *that* and an understanding of *why* a statement is true (or false)" (p. 305). As noted above by Zaslavsky and Knuth (2019), however, there is insufficient work to understand how to aid instructors to help their students use examples more effectively. To begin developing this work on instructor aid, I use the MKT-P framework to understand how instructors of my interview discussed student learning and their own practice.

2.2 Mathematical Knowledge for Teaching

Mathematical knowledge for teaching (MKT) is another well-researched construct. This body of research includes seminal works such as Ball and colleagues (2008), and Shulman (1986). After the development of MKT, researchers began to argue for a different, more specialized type of knowledge denoted mathematical knowledge for teaching proof (MKT-P) (e.g., Lesseig, 2016; Steele & Rogers, 2012; Stylianides & Ball, 2008). Figure 10 below details the MKT-P framework developed by Buchbinder and McCrone (2020).

Figure 10

Mathematical Knowledge for Teaching Proof (Buchbinder & McCrone, 2020)

Type of MKT-P	KLAP: Knowledge of the Logical Aspects of Proof	KCS-P: Knowledge of Content and Students	KCT-P: Knowledge of Content and Teaching
Description	Knowledge of different types of proofs, valid and invalid modes of reasoning, the roles of examples in proving, logical relations and range of definitions and theorems	Knowledge of students' proof-related conceptions, misconceptions and common mistakes	Knowledge of pedagogical practices for supporting students engagement with proof
Related classroom practices	Use precise mathematical language and notation within students' conceptual reach	Identify / anticipate common misconceptions about R&P[(a)] in students' utterances or written work	Identify curriculum opportunities for R&P in diverse mathematical contents
	Identify and correct students' logical mistakes or inaccurate language, and support students' use of correct logical reasoning and language	Facilitate discussions to address common misconceptions about R&P	Design tasks that embed rich opportunities for R&P and make logical aspects of proof explicit
		Make proof concepts explicit and accessible to students' conceptual level	Use productive instructional moves to utilize learning potential of R&P tasks

Like the MKT framework described by Ball and colleagues (2008), most researchers consider MKT-P to be comprised of pedagogical content knowledge (PCK), and subject matter knowledge (SMK) (e.g., Lesseig, 2016; Steele & Rogers, 2012; Stylianides, 2011). For my use of the MKT-P framework, I will focus on the PCK subdomains of knowledge of content and students (KCS), and knowledge of content and teaching (KCT). The two subdomains of KCS and KCT are similar across both MKT and MKT-P. MKT-P gives a more detailed description of what instances of KCS or KCT could look like in a proof setting.

Ball and colleagues (2008) described KCS as knowledge related to understanding the perspectives of their students and intended thinking their students will bring to the classroom. KCS includes topics such as anticipating student answers, interpreting student work, and expected misconceptions held by students in a mathematics context. Buchbinder and McCrone (2020) gave a similar definition described in Figure 10 that has been adapted to a proof specific context. Ball and colleagues described KCT as the knowledge relating to the teaching of mathematics including sequencing and other instructional decisions. KCT includes topics such as how to sequence tasks in a classroom, how to engage students in coursework, and how to develop students as thinkers of mathematics. Buchbinder and McCrone again gave a similar definition for MKT-P adapted for a proof classroom. I use the MKT-P framework described by Buchbinder and McCrone instead of the MKT framework described by Ball and colleagues to have a finer grain analysis tool to use in my interview data.

3. Methodology

3.1 Research Design

This study employs a basic interpretative qualitative research design (Merriam, 2002). As my intention is to study how instructors of ITP courses respond to student work, I was interested in as much context of teaching as possible. Interviewing the participants allowed me to ask them about their classroom and normal teaching practices to gain insight into what their classroom looked like. Further, I was able to ensure participants understood how I was framing example-use in proof to avoid any potential misunderstandings. The interviews were later transcribed and presented in the form of select participants to elicit the variety of responses found in the interviews.

3.2 Participants

Participants for this study include instructors of ITP courses at universities in the Southeastern United States. Emails were sent to the department chair of all accredited universities in the Southeastern region of the United States with a mathematics department to forward to whomever had recently taught an ITP course. For a full description of the demographics and reasoning behind what was chosen, refer to Kirby (2023, chapter 2). Table 8 below describes the varying information of all participants.

Table 8

Participant Descriptions

Participant	Carnegie Classification	PhD Received	Length of Time at University	Private or Public University	Self-described Teaching Style
Dr. Sarah	R1	Mathematics Education	5 Years	Public	Active Learning
Dr. Hubert	R1	Mathematics	30 Years	Public	Lecture
Dr. Phil	R2	Mathematics*	18 Years	Public	Lecture
Dr. Ryan	D/PU	Mathematics Education	20 Years	Private	Lecture
Dr. Tucker	M1	Mathematics	7 Years	Public	Active Learning
Dr. Chris	M1	Mathematics Education	18 Years	Public	Lecture
Dr. Daniel	M1	Mathematics	9 Years	Public	Active Learning
Dr. Nathan	Baccalaureate	Mathematics Education	4 Years	Private	Lecture
Dr. Amanda	Baccalaureate	Mathematics*	6 Years	Private	Active Learning
Dr. Casey	Baccalaureate	Mathematics	19 Years	Private	Lecture
Dr. Joe	Baccalaureate	Mathematics Education	1 Year	Public	Active Learning

Participants denoted with a * under PhD received indicated they also have publications in mathematics education

3.3 Data collection

The participants of this study agreed to a one-hour semi-structured interview online through Zoom. After agreeing to the study, participants were sent an email with the questions on which the student work samples would be focused. The two problems used in this study were titled the *polygon diagonal problem* and the *growing S pattern task.* The growing S pattern task is depicted in Figure 8. Figure 11 provides a screenshot of the polygon diagonal problem in the manner sent to the participants of the study.

Figure 11

Polygon Diagonal Problem

Question 1. Given the number of vertices in a polygon, find a method to calculate the number of diagonals of the polygon. Explain why your answer is correct.

My intention was to not spend a significant amount of time in the interview solving the problems with the participants of the study. I decided to pick these two problems as they were both applicable to an ITP course and I perceived them to be relatively simple problems. Participants of the study were asked if these problems were indicative of the types of questions asked in their own ITP course. Four participants indicated they would use or have used one of the exact questions listed in this interview. Ten of the eleven participants suggested these questions were appropriate for an ITP course and could be easily substituted for a similar question in their own course. The one participant who did not mention whether these questions seemed appropriate for an ITP course described they would instead see questions like this in a discrete mathematics course, which was unique from their university's ITP course.

For each of the tasks in this interview, three student work samples were included to varying degrees of effectiveness. Student work samples were chosen intending to highlight the levels of example-use described by Balacheff (1987). Samples for the growing S pattern task were adapted from students in an ITP course at a large public university in the Southeastern United States. Samples for the polygon diagonals task were adapted from the same task given to students described in Balacheff (1988). Figure 12 shows the alignment of student responses and the intended level of example use described by Balacheff's framework along the continuum I interpret this framework. Figure 9 details each students' response to the growing S pattern task. Figure 13 below describes each students' response to the polygon diagonal problem.

Figure 12

Student Work Continuum and Balacheff (1987) Level

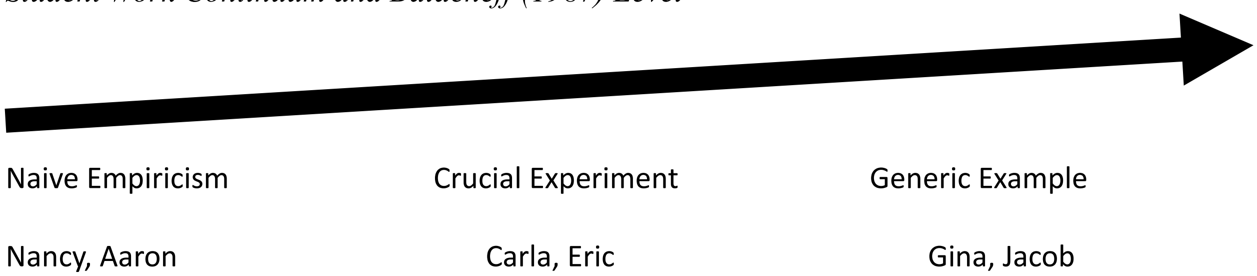


Figure 13a

Aaron Response - Naive empiricism, Polygon Diagonal Task

Aaron

Question 1. Given the number of vertices in a polygon, find a method to calculate the number of diagonals of the polygon. Explain why your answer is correct.

Well, a polygon with 6 vertices has 6 x 3 = 18 diagonals. In a polygon with 8 vertices, there are 8 x 5 = 40 diagonals.

So, the number of diagonals in a polygon equals the number of diagonals from one vertex multiplied by however many vertices there are.

Figure 13b

Eric Response - Crucial Experiment, Polygon Diagonal Task

Eric

Question 1. Given the number of vertices in a polygon, find a method to calculate the number of diagonals of the polygon. Explain why your answer is correct.

Number of Vertices	Number of Diagonals
6	9
7	14
10	?

First, I found that a polygon with six vertices had nine diagonals. Then, a polygon with seven vertices had five more diagonals than that one for a total of fourteen diagonals. This led me to a pattern.

Then, I tried to test the pattern by jumping ahead to a big shape with ten vertices. If my pattern holds, this one should have 35 diagonals. When I tested this by drawing it out, the pattern held. This means the claim is true.

Figure 13c

Jacob Response - Generic Example, Polygon Diagonal Task

Jacob

Question 1. Given the number of vertices in a polygon, find a method to calculate the number of diagonals of the polygon. Explain why your answer is correct.

I started from a polygon with 7 vertices. The number of diagonals to each vertex is n – 3.

For a polygon with 7 vertices, each point has two neighbors. So:

7 – (2 + 1) The 7 is the number of vertices, the 2 is the number of neighbors, and the 1 is the point itself

In general, n – (2 neighbors + itself)

Then take this and multiply by the number of vertices. This counts each diagonal twice, so divide by 2 for the answer.

So, my final answer is n(n-3) all divided by 2.

The participants of this study were shown either the growing S pattern task or the polygon diagonal task first. The three student work samples within each task were always shown consecutively. This order was also randomized for each interview with no two interviews having the same sequence of student work shown. The polygon diagonal task was shown first in 6 interviews while the growing S pattern task was shown first in 5 interviews.

The first part of the interview involved participants engaging with the student work samples of either the polygon diagonal problem or the growing S pattern task. For each sample of student work, the participants were asked to talk aloud and mention anything they noticed about the student's solution or any questions they may have about the student work. These noticings typically included remarks about expected student misconceptions or strategies as well as how the participant would respond to a student providing this work in their classroom. The context given for the problems was that these six students were enrolled in the participant's ITP course and submitted these assignments as homework to be graded. Participants were asked how they handled homework in their classes. Four participants mentioned the use of final submission portfolios where students were allowed repeated submissions of homework assignments. Students at the end of the semester would choose a select number of their submissions from across the semester and put these together for a portfolio which consisted of a majority of the homework grade for these instructors. The other seven participants all mentioned they used standard homework practices including assigning problems from a course textbook or class notes to be collected later. All participants used a 100-point grading scale for homework assignments that ranged between 10 and 20 percent of the final grade average for students in their class.

The participants of this study responded to the six student work samples for the first part of this interview. The rest of the data for this study comes from the final part of the interview where participants were asked to rank the student work samples in order of the best attempt at a proof or most progress toward proving the claim provided. A participant ranking a student solution a "1" meant they believed this solution to be the best attempt or most complete attempt at proving the claim. After ranking these student solutions, participants were asked to give explicit instruction or questions for the students ranked first and last in their individual rankings. This instruction could be specific questions the teacher would ask to gain further insight into the student's work or related assignments or problems to help move the student forward in their thinking.

3.4 Analysis

The primary data source for this study comes in the form of transcriptions of the interviews. Transcription was done by talk turns to allow for more context in what was spoken. I define a talk turn as all utterances from one participant of the interview until they were interrupted (by either a statement or question) from the other participant. I did not code any talk turns taken by the interviewer. I report on excerpts from the transcripts for two parts of my study: the first task of responding to student work, and the final task of responding to the student with explicit instruction to help move the students forward.

Before any coding of the data, I had three meetings with a research team I frequently work with. In the first meeting, we as a research team came to agreement on instances of what would be considered instances worth coding using the MKT-P framework. We then coded a small section of an interview using two of the categories listed as types of MKT-P under Figure 10: knowledge of content and students (KCS), and knowledge of content and teaching (KCT). After this initial coding section, we met together to discuss all lines until we reached agreement.

In our second meeting, we further differentiated what should be considered KCS and KCT, and if a sentence should be allowed to be coded for multiple instances of KCS/KCT. Our guidelines for coding KCS and KCT included the instances described in Figure 10 under either the related classroom practices or the description of the category. For instance, Figure 14 shows a short excerpt from an interview with Dr. Nathan that was coded multiple times. Instances of KCS were coded in red, while instances of KCT were coded in blue. The section below coded in red was chosen as KCS due to the relation of explicitly discussing the description of KCS described by Buchbinder and McCrone (2020). The section below coded in blue was selected as KCT due to the discussion surrounding an instructional move or a related task by introducing inductive thinking to the student.

Figure 14

Instance of Coding Dr. Nathan's Transcript (from Gina's Work)

> But otherwise, again, perfect reasoning and everything. And certainly I could, if it's just a thing of I'm trying to give her ideas of other approaches or something that could be something to do too think about how either, you could think about this using induction or potentially, like one way I thought about doing this too, was to really kind of see where the n squared comes from. You could even just move some of this around and make it more of a square kind of think about this in a more kind of dynamic way with that picture.

We later discussed if a line could be interpreted as both KCS and KCT if this should be a separate code or allowed to be double coded. Figure 15 shows an instance (coded in purple) of what was determined to be both KCS and KCT depending on the interpretation of the coder and amount of context included from the conversation.

Figure 15

Instance of Multiple Coding from Dr. Ryan (from Aaron's Work)

> But that's not our setting here. Our setting is we're in and out. We're in this kind of class, this is your peers, this is what we're doing. You know, let's own that let's claim it. Let's, let's use our notation our language or vocabulary to full effect. And so this one doesn't do that at all actually.

In this Figure, Dr. Ryan was previously discussing aligning the submitted argument from students to the intended audience of the reader. Dr. Ryan described the differences between describing a mathematical statement to a friend or parent unfamiliar with mathematical proof versus someone intimately familiar with proving. The sentences coded in red (KCS) were decided to highlight Dr. Ryan describing his knowledge of the classroom setting and what the audience for his students would be. To make understanding proof accessible to the students, Dr. Ryan discussed the idea of the peers of the student writing the proof relating to one of the related classroom practices described in Figure 10. In the last sentence, Dr. Ryan remarks about Aaron's work that he does not think Aaron accurately presents his argument to the intended audience. This relates to the description of the category of KCS given by Buchbinder and McCrone (2020). I interpret this as what Buchbinder and McCrone defined as "knowledge of proof related conceptions." Dr. Ryan understands that what Aaron has argued is not using the language or vocabulary developed in his proof class to the full effect.

The section coded in purple was coded as his knowledge about students' audience and the language used by his students when arguing. For a similar reason to what was coded as KCS in the previous sentence, this section was coded following the categories described in Figure 10. Further, it was also decided that these two sentences could be interpreted as a move Dr. Ryan would make in his classroom as an instructor and explicitly ask students to consider the notation and vocabulary to full effect. A mention of an instructional move or related task was decided to be coded as KCT following the related instructional practice described in Figure 10.

In the final meeting, we transcribed a section of another transcript from Dr. Sarah until we agreed on the broad categories of KCS, KCT, or both. All talk turns that were coded in any way were later sorted by participant to get a list of all instances of KCS, KCT, or both codes for all participants. To answer my research question for this manuscript, I then took instances of KCT and looked for the different ways instructors responded to the students in both parts of the task being analyzed. I first listed all the ways in which instructors gave recommended instruction or questions to any student at any part of the interview. I originally anticipated the codes to be the subcategories of KCT described in Figure 7 along with the description of Figure 10. However, I stayed open to the possibility of new codes outside the scope of the categories described in Figure 10 should a section of the transcript not fit neatly into one of the categories. This could arise from either feeling the need to fit a code into two of the categories described in Figure 10 or what was deemed a clear instance of KCT that did not fit in any of the categories of Figure 10. I used these codes to make a list of how each participant responded to student work samples.

Once I had a list of categories of how each instructor responded to student work at any point in the interview, I then checked for similarities in responses across the interviews. For this study, I will focus on three instructors and their responses in the interview task. These instructors were chosen to represent as broad a scope as possible from my selection criteria as well as for the depth of their responses. The participants span the three broad Carnegie classifications, professorial ranks, teaching style, public or private universities, and type of Ph.D. earned. Further, the three participants included in these findings represent the full spectrum of responses to the type of interview questions I asked. These three participants gave some of the most in-depth responses with clear reasoning for their positions throughout the interview.

4. Results

I will answer my research question, In what ways do instructors respond to student work samples involving examples in an ITP course, by reporting on how three participants of my study responded to student work samples across the naive empiricism, crucial experiment, and generic example level of understanding according to the continuum idea I discussed aligning with Balacheff's (1987) example-use framework. I will report primarily on data from the first part of the interview task where participants responded to samples of student work, as well as the final part of the interview task where participants were explicitly asked to reflect on how to respond to student work with instructional moves. Each of these sections will include a story depicting how different participants of my study responded to the varying uses of examples shown to them in the student work.

These three participants were chosen to demonstrate the variety of responses received over the course of the tasks in the interview. This variety includes which task the participant was shown first in the interview, how the students were ranked in the ranking task, and common instructional strategies recommended to students such as suggesting students work on alternative tasks or how formal the language in a proof should be. The three participants included in this study expressed their thinking clearly and gave the most detailed answers to questions asked in the interview. Further, these three participants span the collected demographics with participants from all clusters of Carnegie classification, varying types of Ph.D. received, self-reported teaching style, public or private university, and length of time teaching at the university level.

The first participant's comments I will share come from Dr. Daniel. Dr. Daniel is an associate professor at a public M1 university with a Ph.D. in mathematics. Dr. Daniel expressed to me that he primarily uses active learning style lessons in the form of small group work and student presentations in his class. Dr. Daniel was shown the growing S pattern task first in the interview before the polygon diagonal problem. The order of the six student work samples he saw was: Carla, Gina, Nancy, Jacob, Eric, Aaron. In the ranking task at the end of the interview, Dr. Daniel ranked the students in the order (from 1 to 6): Gina (generic example), Jacob (generic example), Carla (crucial experiment), Aaron (naive empiricism), Nancy (naive empiricism), Eric (crucial experiment). Recall the ranking task had participants rank the student they thought had the best written product of a proof as a 1, and the student with the most work needed toward a proof as a 6.

The second participant's comments I will share come from Dr. Amanda. Dr. Amanda is an assistant professor with a Ph.D. in mathematics who was submitting an application for tenure at the time of our interview at a private baccalaureate college. Dr. Amanda remarked she had also published in mathematics education journals since completing her Ph.D. Dr. Amanda expressed to me that she uses more active learning style lessons in her class than lecture classes, although she uses both throughout the semester. She defined active learning style classes as getting her students engaged with the learning process through group work and class discussions. Dr. Amanda was shown the growing S pattern task first in the interview before the polygon diagonal problem. The order of the six student work samples she was shown are: Gina, Carla, Nancy, Aaron, Eric, Jacob. For the ranking task at the end of the interview, Dr. Amanda ranked the students in the order (from 1 to 6): Gina, Jacob, Carla, Aaron, Eric, Nancy.

The third participant's comments I will share come from Dr. Ryan. Dr. Ryan is a full professor at a private D/PU university with a Ph.D. in mathematics education. Dr. Ryan also has experience as a high school mathematics teacher for 10 years. Dr. Ryan expressed to me that he uses primarily lectures in his classroom, although he sprinkles in some active learning style activities in the form of student presentations a few times a semester. Dr. Ryan was shown the polygon diagonals task first in the interview before the growing S pattern task. The order of the six student work samples he was shown are: Eric, Jacob, Aaron, Nancy, Gina, Carla. In the ranking task at the end of the interview, Dr. Ryan ranked the students in the order (from 1 to 6): Gina, Carla, Jacob, Nancy, Eric, Aaron.

4.1 Naive Empiricism

Recall the two student work samples highlighting a level of example-use described as naive empiricism: Nancy and Aaron. When responding to student work, either Nancy or Aaron was ranked last by 8 of the 11 participants of the study. The most common response when initially reading through Nancy or Aaron's work was the idea of the student using a proof by example. For instance, Dr. Daniel remarked about Nancy, "So of course, the trouble is, this is a very common issue in intro to proof class, where working a few examples are seen as kind of like justifying a general trend." This comment or something similar about a hesitation to prove by example was made by all 11 participants of the study. These remarks were built upon how the participants of the study suggested they would help either Nancy or Aaron in their class.

Dr. Daniel started discussing Nancy's work wondering if she had done more work that was not submitted. Dr. Daniel opened his comments about Nancy's work with:

> Yeah, see… here I would suspect that there were pictures in the background that she didn't submit. And at least for what she wrote, I would really want to see them because like, we kind of know that 10 is 3 squared plus 1 and then 4 squared plus one… So yeah I have trouble with her answer to question 3. There is also some notational weirdness I would say. (Referring to her work) What is that? Is it like the real numbers? Yeah that just seems distracting.

Dr. Daniel continued to express concern for the work Nancy had shown throughout the interview. Following the above comments about wondering where Nancy started her reasoning, Dr. Daniel drew a connection between the work of Nancy and the work of Carla (which he had seen prior to commenting on Nancy's work):

> What she means is that the number of squares is n squared plus one for the nth step. Yeah, and this is kind of going back to the like, second slide's (Carla) mistake of just saying like, Oh, I've listed all these cases. Therefore the formula is true. Which shouldn't be convincing to people.

Like his comments on Nancy, Dr. Daniel held issue with the work presented by Aaron. After spending a few minutes working through understanding what Aaron wrote down, Dr. Daniel commented, "I would like them to be a little more precise about the formula. Just so I don't have to think really hard about what they mean by number of diagonals from one vertex." This desire for precision continued throughout the interview as Dr. Daniel commented on the students he would later rank near the bottom of the student solutions in the ranking task.

Dr. Daniel held issues repeatedly in the interview with the way students tended to use examples to draw general conclusions. His idea of students trying to prove by example was repeated numerous times throughout his comments. This led to some of his recommendations to the struggling students. After commenting for a while about the work Nancy had shown him, in the first part of the interview Dr. Daniel commented unprompted by me about how he would help combat the problem of proof by example, "I just say something ridiculous. Like, you know, suppose you think every number is less than a billion. It'll take a while to come up with an example that violates that if you go write numbers one by one." Dr. Daniel remarked he used strategies like this repeatedly in his ITP course when students were confused why checking a few examples were insufficient to prove a claim.

Dr. Amanda had similar remarks as Dr. Daniel when discussing Nancy's work. Dr. Amanda first started with some positive thoughts about what Nancy had done,

> They're using n squares and stuff and it seems like it's totally away from the geometry. I like the definition, or like how these formulas came to be. Whereas like, the very first example (Gina) was very like, based off of the actual structure I have…This one (Nancy) is much more, I found a pattern let's make sure it works. Although then they're not sort of verifying like the last example (Carla) was kind of in between where they had here is what the picture would look like. Here is what the map looks like.

Similar again to the comments from Dr. Daniel, Dr. Amanda held issue with the idea of verification from Nancy and the possibility of work that was excluded from the final product.

> How are you verifying that it is correct? Other than sort of doing, saying 101 (Referring to plugging in the number 10 in the formula and getting a value of 101) How do you know that? That is right. Maybe they had other unwritten work that verified it. But it looks like they have sort of moved away from… I don't know, I would personally be worried about this category of correctness.

Further when commenting about Nancy's work, Dr. Amanda commented on an instructional strategy she would use in her class, "I would still sort of make a comment like, well but how do you know? What is your conclusion?" She gave a similar response when commenting on Aaron's work, "I think my initial reaction to this would be to ask questions for this (referring to how Aaron is convinced of his answer) exactly. Those details, at least for an intro class, I would really want them to sort of spell it out." Dr. Amanda was consistently adamant about wanting students to show their reasoning, particularly for students that were ranked lower in the ranking task at the end of the interview.

At the end of the interview after ranking Nancy sixth, Dr. Amanda offered her comments about what instructional strategies or moves she would make if Nancy were a student in her class.

> I just have a follow up question with the correctness thing. Where okay, so it works for 10. But what if the number we needed to check was like, a million or something that is so obnoxiously large that they couldn't… you know… you could do the number but you couldn't attack it with a picture. So then how do you sort of, pushing them to think abstractly because it is too large to sit down and try to calculate.

Again, we see a common instructional strategy of going to the extremes with students at the naive empiricism level of example use to try and break the habit of doing a proof by example. Dr. Amanda goes further than Dr. Daniel and mentions the difficulty of drawing a picture for very large numbers.

When responding to the students selected to exhibit naive empiricism, 8 of the 11 participants had a similar instructional strategy described above by Drs. Daniel and Amanda. These instructional strategies varied from presenting extreme cases intended to be impossible to draw, or to use large numbers to make calculations seem impractical. These extreme cases were similar to what Stylianides and Stylianides (2009) described as "Monstrous Counter-examples" (p. 327) and involved the instructors pushing their students to think of cases that may provide conflict with their initial reasoning. Further, the notion of wanting precision in work and the possibility of student work being left out was mentioned by 10 of the 11 participants.

Dr. Ryan had similar ideas of precision and potentially work being left out as Drs. Amanda and Daniel. Dr. Ryan did not mention an interest in introducing extreme examples to break the students' conceptions about naive empiricism. Rather, Dr. Ryan focused on students using examples more familiar to them. Dr. Ryan mentioned an idea he repeated frequently about alerting students to the audience of their work. Dr. Ryan began his interview by talking about audience:

> So one of the points I make to students on this, so this would be a good opportunity, you know, who is the audience for your work here? And a common mistake I will point to them is, if you assume I'm your audience, which is a natural assumption, but it is not actually a good assumption. It's, you may well leave too many things out. Thinking that I will know what you mean. And that is problematic. I suggest early and often to them that, so who is your audience? And they realize well, it isn't you is it? They will then say it is a peer. And I will say, OK. Yeah someone else in the class may be even better. And this seems to work with them pretty well. But how about a peer is you a week from now? A month from now?

This notion of considering an audience was weaved through the whole interview with Dr. Ryan and was mentioned in all six student responses. Regarding Nancy and Aaron, Dr. Ryan first focused on the audience and a similar notion of precision that Dr. Daniel had mentioned. When responding to Aaron, Dr. Ryan commented:

> I did two examples and so now it is true in general (referring to the work done by Aaron). But without naming, without putting a variable expression. So to me

> again, this is a ripe time to talk about audience. If you are talking to your mom or dad or grandma at home about, "Hey there is this cool problem I am doing." Okay sure, you maybe keep it in prose and I may not want to see an algebraic expression. But that is not our setting here. Our setting is we are in and out. We are in this kind of class. These are your peers, this is what we are doing. You know, let's own that. Let's claim it. Let's use our notation and language or vocabulary to full effect. And this one doesn't do that at all actually.

When Dr. Ryan was responding to Nancy, he was interested in seeing more work and reasoning from what Nancy presented.

> I know these weren't asked for in the question, but admittedly I guess I would be interested in some insight from her as to well, how did you come across this? Was it a trial and error? And oh look, there is one more than a perfect square. Was there something else going on? If I was having a conversation with her about if she came and asked me a question, I would be curious to know. That looks nice. How did you get that idea?

For the ranking task, Dr. Ryan ranked Aaron in sixth place. When commenting about what he would do if Aaron were a student in his class:

> For Aaron's, it would probably be helpful if I had an example. Either something we had worked on in class already. Ideally it would be one that wasn't cold to him or an exercise done in the book or something like that. Next best, if I had one kind of ready at hand that I can offer to him and talk him through that would be harder because I would be asking him to entertain a new problem while I am trying to help him understand this problem. I probably have him talk out loud to me. So tell me what are you thinking here? How did you notice that? Good observation. Is there any way you could say that more generally? You know kind of help him build some iterative thinking. You know, if I'm seeing that Aaron's having trouble more generally, are you part of a study group? Encourage collaboration with a peer.

Overall, the participants of this study recommended students at the naive empiricism level of understanding example-use to consider alternative examples to help build intuition to begin to work towards a more formal proof. These ideas ranged from having students use examples with purposefully large numbers, to examples that are difficult to draw. Dr. Ryan mentioned introducing examples that are like what students are already familiar with and instead focus on having his students explain what was happening in the problem to investigate their understanding. By having students use examples in this way, participants seemed to be encouraging the students to advance from using examples in a naive empiricism level of example-use to a more abstract or purposeful approach to examples. This push may have students reach the crucial experiment level of example-use or the generic example level of example-use.

4.2 Crucial experiment

The two student work samples aligned with a crucial experiment level of example-use were Carla and Eric. Carla was ranked third in the ranking task by 9 of the 11 participants. Carla was ranked second once and fourth once. Eric was ranked fifth by 8 of the 11 participants. He was ranked sixth by the other 3 participants. It is important to note a common theme when responding to Eric was the focus on some of the comments made by Eric including, “This led me to a pattern.” and, “If my pattern holds.” All participants expressed concern for these phrases and ranked Eric lower than at least one of the students at the naive empiricism level based on the understood lack of clarity presented by Eric’s argument. For this manuscript, I will focus on comments made about Eric’s work regarding his use of examples in the crucial experiment level of example-use.

As a final note for this section, since most participants ranked the two student work samples aligned with a crucial experiment level of example-use between 2 and 5, the final part of the interview where participants were asked to provide explicit instructional moves to help students move forward largely includes responses to students at the generic example or naive empiricism level of example-use. Thus, the comments included here, as related to crucial experiment, were primarily unsolicited remarks for instructional strategies provided by participants in the first part of the interview when responding to student work.[1]

Dr. Daniel's comments about Eric were almost entirely focused on the statements mentioned above and are outside the scope of this manuscript. Dr. Daniel had high remarks for the work shown by Carla. When responding to Carla's work,

> I think the picture is actually pretty convincing. And it might even be something like a proof by pictures book would include. So like, it's not terrible. I guess I am kind of tripping over her statement I tried the smallest case and a larger case. Since the formula holds for both, this makes sense for all. Because that can't be like a generic principle. If you check two cases, that like verifies a formula. So it is not uncommon.

Dr. Daniel gave some unprompted instructional advice to Carla when responding to her work. He said, "I would try to push her towards either thinking through like, how this picture that she drew can kind of be generalized." This sort of remark for a push towards generalization was mentioned by 6 of the 11 participants when responding to students at the crucial experiment level of example-use. The remarks given by participants seem to push students at the crucial experiment level of example-use to consider how to argue from a generic example level.

Dr. Amanda similarly had some positive comments about Carla's work with how she approached examples. Dr. Amanda noted, "It looks like they are sort of thinking beyond can I verify it for the examples we have been given? This was really interesting when they came up with the idea of a base case." An idea of inductive reasoning with the base case was noted by 7 of the 11 participants when referring to Carla's work.

[1] Participants of this study were asked to reflect on what they would say to the students ranked first and last in the ranking task for the interview. As most of the students in the crucial experiment stage were ranked somewhere in the middle between 2-5, I have limited data available for this stage.

When referring to Eric, Dr. Amanda focused primarily on the wording used in Eric's argument. This was compounded with how Dr. Amanda interpreted the choice of numbers Eric used in his proof, "So then when they are jumping ahead, are they just adding? It is unclear to me what they are adding each step." Later in the interview, Dr. Amanda came back to Eric's work and again took issue with how Eric was choosing his numbers to use, "There is some pattern he is seeing there. And then he is jumping ahead. It is not sort of grounded in anything other than numbers."

In the final part of the interview, Dr. Amanda commented on Carla's work again and what she liked about Carla's submission:

> Carla is demonstrating this pattern on a particular example. So she definitely has an idea of what is going on. Maybe unverified like Aaron and Eric, but there is a hint that there is more. There is a hint that there is the n plus 1 n minus 1. There is a pattern that exists on actual structure that can then be generalized.

These comments are promising as this helps shed light on how Dr. Amanda intuitively values the work of a student at the crucial experiment stage of example-use above that of the student work at the naive empiricism level of example-use. As a reminder, no participants of this study were made aware of the frameworks used in this study by Balacheff (1987) or informed which student work samples were at each level of example-use.

Dr. Ryan had similar comments about generalizing the work when responding to Eric's work.

> It is a major issue for a sizable number of students to realize that I am going back to the logic of it is essentially, you are stating something generally about a polygon with any number of sides. How do you do that? How do you take your good work here and generalize it? I guess it depends from the instructor standpoint, where are you in the course? What are you asking from them and expecting of them? If you are expecting something real formal, there is some

> critique to offer about, you know, being careful about the use of a specific number of vertices. A strong part of this is that he starts with the generalization.

In addition to the ideas of considering the audience that Dr. Ryan mentioned previously, he had further insight into the work shown by Carla about how to make an argument.

> I think there is a way to frame and connect and tie it together that would be more effective. This might be an example of something where I would, and some students really struggle with this and push back on it, but I would make a comment on how… Ok things here are correct. How do we make an argument? I use the language of argument a good bit in class. How do we make an argument that what you have here, tie it together and have a nice flow to it. Use sentences. Again some of this might depend on at this point are you looking for that kind of formality in your class or not. Even if not, this case of the continuum we are on here exists. We are not on a binary of yes it is right. No it is not right.

This quote highlights the instructional strategy Dr. Ryan suggested for both Carla and Eric.

> What you have here is correct. But can it be expressed in a form and flow that is more effective and stronger? How do you make your case for what you are observing and concluding stronger? So that I would find a real opportunity here that didn't quite exist in the others? Because they (Carla and Eric) were missing some kind of more fundamental pieces.

For the work shown by Carla and Eric, there was limited data to discuss instructional strategies. Most of the participants of the study had similar thoughts about the student work indicated by the consistent ranking of Carla in third place and Eric in fifth place. The ranking of students at the crucial experiment stage, particularly Carla, shows instructors' intuition about student work aligning with the continuum of example-use described by Balacheff (1987). Further, instructors began to remark on the type of examples used in a description akin to how Balacheff defined crucial experiment. Dr. Amanda noted, "... there is a hint of n plus 1 n minus 1." and a push to generality.

However, the focus on Eric's language in the argument detracted from a discussion about the use of examples in his argument. Further, the tendency to rank these participants in the middle excluded any discussion of explicit instructional intervention for students at the crucial experiment stage of example use in my study by most of the participants. Those who ranked Eric last focused in their instructional interventions on focusing on Eric to explain his reasoning more and not rely on comments like, "I saw a pattern."

4.3 Generic example

Gina and Jacob are the last two student work samples and are aligned with a generic example level of example-use. Gina was ranked first by 10 of the 11 participants in the ranking task. Gina was ranked second once. Jacob was ranked first once, second nine times, and third once in the ranking task. Universally, participants were the most pleased with the students at the generic example level of example-use and made frequent comments about these students compared to the others. For instance, Dr. Amanda remarked about Gina comparing to the other student work samples, "Gina's is like, it (her work) doesn't rely on the size at all. I mean she relies on n."

Dr. Daniel had high praise for both Gina and Jacob. When commenting on Jacob's work at the beginning of the interview, "I mean, it seems pretty convincing. Like at the intro level, even maybe at a more advanced level. I'm not sure I could even come up with an objection." As Dr. Daniel continued thinking through Jacob's work throughout the interview, he started noting slight formatting points in Jacob's work.

> I would say there is like a stylistic issue that makes it a little bit muddled in that like he starts talking about seven vertices and then quickly shifts to n. And so it can, I think of introductory classes as a little bit of a rhetoric class, where you

> want to make like an argument clear and convincing for someone else. And so if you want to start by explaining a specific example, that can be useful, but if you want to talk about the general case, just talk about the general case. So there are some style issues I would say. But I think the core of the argument is good.

Dr. Daniel had similar comments when reading Gina's work. "Some people get nervous with proof by pictures you know?" The idea of worrying about the formatting and specifics of the argument became apparent with the participants when responding to student work samples at the generic example level.

At the end of the interview, Dr. Daniel offered some feedback to give to Gina as he had ranked Gina first in the ranking task.

> It can be a helpful intellectual exercise, even though I am convinced by her picture honestly. But it can be a helpful exercise to see if it can be proven by induction. Or at least like ok you know the answer. If you have convinced me you have convinced yourself that this is the answer. Is there any other way you can see how to argue for the same thing?

This idea of finding a new method to do a proof another way was mentioned explicitly by 7 of the 11 participants. The remaining 4 participants all mentioned an idea about formalizing language in an argument without mention of an alternative proof for the question. Dr. Daniel gave instructional responses for both above common categories. Dr. Amanda gave a response focusing instead primarily on wording.

Dr. Amanda had a similar structure for her responses to that of Dr. Daniel. Dr. Amanda started off with high praise for both Gina and Jacob. When commenting on Gina's work at the beginning of the interview, "It seems like they have nailed the pattern here. I would consider this a really strong answer, especially since they have got sort of, gotten from the example to a more general description that can be operated on and those kinds of things." Dr. Amanda later referred to Gina multiple times in comparing her work with that of Nancy and Carla. Dr. Amanda used the wording of ‘championing’ a student’s work when she really liked something they did and thought it would be worth introducing to her class. When comparing the work of Nancy to that of Gina, “I would be like you are a champion (referring to Gina)! Look at what you have done. You may not think it is great, but it is really really great that you have checked your formula and the existing ones.”

Similarly, when Dr. Amanda was responding to the work of Jacob, she compared his answer to that of Aaron and Eric. “Okay, so very fabulous. That was like all the details. Like this is formalizing the one (Eric) and then correcting the issue with double counting of the other one (Aaron).”

Later in the interview, Dr. Amanda had a similar progression to Dr. Daniel of finding some work to be done with the formatting of the student responses at the generic example level. When commenting on Gina’s work after the above comments had been made, “I think I would still sort of make a comment like, how do you know this continued and what is your conclusion?” In reference to Jacob’s work, “Jacob’s solution is in a proof class, or sorry, a writing class. I might make some comments about the structure of the writing itself. There is all of the sort of content there, but maybe making sure that it is sort of written in complete sentences and that sort of thing.”

In the ranking task, Dr. Amanda ranked Gina first and Jacob second. She offered her thoughts and recommendations to both students. When referring to Jacob’s work, “I would probably ask him something simple like, What if you had started with eight vertices? What would have happened? Can you extend your argument without relying on the number of vertices that you start with?” In this quote, Dr. Amanda takes issue with the example used by Jacob and does not believe it should be included in the final product. She mentions this is why she likes Gina's argument more, "He (Jacob) uses an example, which makes Gina's argument better than Jacob's argument in terms of extending it. The actual argument doesn't rely on the seven vertices. I don't know if he sees that."

Dr. Amanda offered this advice in addition to the advice on formalizing the work of Jacob, "I would push him to sort of argue without saying I started from seven blah, blah, blah to make sure this is really cemented in." These types of comments were shared by 4 of the participants.

Dr. Ryan again praised the work of both Gina and Jacob. When commenting on Gina's work, "She generalizes it. I would say you know, I applaud the nice proof. It is labeled and it's… I think she has identified everything pretty well." Also, just like Drs. Daniel and Amanda, Dr. Ryan followed his initial praise of the generic example student work with some issues of formality, "Maybe just a little bit of more framing in her paragraph of explanation. I mean it is very good. A little bit of emphasis of the conclusion. Again, sort of depends on what or how formal you want something here. So I mean, that kind of context I think is always relevant."

In the ranking task, Dr. Ryan ranked Gina's work first. When asked to give advice for what he would do as an instructor with Gina in his class, he said,

> I would offer affirmation. For Gina, I would affirm that she labeled and has a nice depiction. She was pretty clear in that regard. And so she made the generalization, the abstraction, that is typically pretty critical in a proof. So I would want to make sure she hears that. And probably using again the language of argumentation, maybe ask her who is the audience you know? Kind of nudging her to building a flow to a set up development or conclusion. Let's figure out a narrative that lets you persuade somebody that, you know, beyond a shadow of a doubt what you are saying is true.

For students at the generic example level of example-use, participants typically responded with strong affirmation of the proving attempts given by students. The feedback to students was often focused on finer details of proving such as the wording or formality of the proof as opposed to the larger logic or foundations as was seen in the naive empiricism level. Common instructional strategies suggested to students were to either try to prove the problem in an alternative method, or to try and rewrite the problem in a more formalized manner.

5. Conclusion and Discussion

In this study, I sought to characterize how instructors of ITP classes responded to student sample work with varying uses of examples to aid their proofs. My research question was: In what ways do instructors respond to student work samples involving examples in an ITP course? I answered this question by separating the responses of my participants of this study by the example-use framework described by Balacheff (1987).

A naive empiricism level of example-use was illustrated through the work of Nancy and Aaron through their use of a few examples and their subsequent jump to claiming a general claim was true in all cases. Instructors' pedagogical suggestions for helping students who were at the naive empiricism level typically involved engaging students in something that looks more like a crucial experiment (i.e., testing varied or extreme cases). I find it interesting no participants expressed a desire for either Nancy or Aaron to focus instead on formatting the argument to a more formalized writing style typical of textbook proofs; such comments were left for students deemed more advanced.

Carla and Eric represented students at the crucial experiment stage of example-use, which consists of students using purposefully selected examples to try and show that the claim holds in a variety of extreme or creative cases, and therefore would hold in all cases. For students at the crucial experiment level of example-use, limited data in this study clouds common instructional responses. However, the participants who shared their thoughts about the student work demonstrated an overlap between the instructional responses they would recommend for students at both the naive empiricism and generic example levels of example use. These instructional responses included a desire for more formalization in the written proof and the consideration of other examples to gain intuition for the problem. We see through the comments of Dr. Amanda that instructors intuitively understand the spectrum of student responses aligning with Balacheff's (1987) example-use framework. Dr. Amanda commented to Carla, "there is a hint of n plus 1 n minus 1." and discussed a tendency toward generality. With this implicit understanding, the participants of the study consistently wanted students to move forward in levels of example-use from one level to the next. For the instance of crucial experiment, instructors wanted to push the students to the generic example level of example-use.

Gina and Jacob were the students selected to exhibit a generic example level of example-use. Generic examples are characterized by using an example in a manner through which one can build intuition and craft an argument distinct from the specifics of the example used by the prover. Instructional responses for students at the generic example level of example-use included looking for alternative approaches to prove a problem in another way or a push for higher formality in the written argument. This represents the primary time in the interview when participants expressed a desire for students to move out of the use of examples and into the formality more common in textbook mathematics proofs. The goal of instructors moving their students along continues here as instructors have finished moving their students along the continuum of example-use described by Balacheff (1987) and now help their students take the next step toward formalized proving.

It is important to note here again that no participants of this study were made aware of the example-use frameworks described by Balacheff (1987) or any related frameworks. Further, participants were not told there was a hierarchy of student work samples to be shown, or what stage of example-use each student was at. Despite this, instructors still made decisions and ranked student work largely aligning with the example-use framework described by Balacheff. After moving through the levels of example-use, instructors then felt confident to suggest students move toward formality in their proof.

As the field of example-use in proof continues to research how to aid practitioners of ITP courses with instructional interventions or tasks, it is important to consider what instructors of these courses are currently doing to help their students learn to prove. The findings in this study suggest that instructors of ITP courses largely have an intuition about the nature of the continuum of example-use. Further, these instructors currently anticipate helping students move along this continuum one step at a time. Rather than design an instructional intervention to help students at the naive empiricism level of example-use develop an understanding at the generic example level or even beyond to a formalized proof devoid of an example, mathematics education researchers should consider aiding instructors by having the students develop one stage at a time. This would align with recommendations for professional development argued by Desimone and

Garet (2015). However, not all instructors see the merit of examples in the written work of students.

Kirby (2023, Chapter 2) found that many instructors may be hesitant to have their students include examples in their formal written work at all. This was further seen in the comments made by the participants in this study particularly at the generic example level. I recommend first approaching instructors of ITP courses with activities and interventions built on the methods described by the participants of this study.

This study has two main limitations. The first limitation of this work is the findings for the crucial experiment level of example-use. As most participants did not rank students at the crucial experiment stage first or last, there is limited data to answer how the instructors respond to student work. Future work should investigate this level more closely to find any differences between students at this stage and students at either naive empiricism or generic example levels of example-use. Second, the type of question and proof used likely affects how instructors respond to students. It would be interesting to compare instructor responses across proof types including induction, indirect proofs, and direct proofs. Future work should also consider having participants give their recommendations across proof type as well as varying the level of student work shown.

Article 3: Practitioner Work (Intended for PRIMUS)

Engaging students in example-use in proof

1. Introduction

Students often have difficulty learning to prove (i.e., Stylianides & Stylianides, 2009; Weber, 2010). Often, students are learning to prove for perhaps the first time in an introduction to proof (ITP) course (Boyle et al., 2015). One strategy common to students as they learn to prove is the use of examples to help build intuition for their proofs (e.g., Balacheff, 1988, Buchbinder & Zaslavsky, 2019; Healy & Hoyles, 2000). However, students often misunderstand the purpose of examples or try to generalize an argument based on a few examples (e.g., Balacheff, 1987; Harel & Sowder, 1998, 2007; Healy & Hoyles, 2000). When used productively, examples can help students gain intuition for how to approach proofs (Aricha-Metzer & Zaslavsky, 2019).

Aricha-Metzer and Zaslavsky (2019) argued, "In this sense, a generic example can be seen as a bridge for students, as it may help them move from empirical views and misleading intuitions to a justifiable conviction *that* and an understanding of *why* a statement is true (or false)" (p. 305)." Recent shifts in mathematics education research have begun to investigate how to help instructors of ITP courses get their students to use examples more productively (Aricha-Metzer & Zaslavsky, 2019; Kirby, 2023, chapter 2,3; Zaslavsky & Knuth, 2019). This article presents a lesson to be implemented early in the semester for an ITP course to help students understand how to use examples more productively in their proving. The lesson presented in this article is developed in consideration with the findings from chapters 2 and 3 of Kirby (2023) dissertation work.

The goal of this lesson is to implement a research to practice lesson through the findings of the author's dissertation work.

2. Theoretical Framework

2.1 Example-use

To understand the role of examples in a proof setting, I use the example-use framework described by Balacheff (1987). Balacheff described 3 levels of how students commonly understand examples: naive empiricism, crucial experiment, and generic example. I interpret these three levels to be along a spectrum of understanding that students will advance on as they gain experience in proving.

Students may exhibit naive empiricism when they try to generalize an argument based on a few examples. Often, this may be the typical case of a 'proof by example' that is so common amongst students. The *proof* students may try to argue comes from the assertion that if something holds true for as many cases as you want to check, surely it is true in general.

A crucial experiment level of example-use is a slight step to the right on the spectrum of understanding where students now start to consider the specific examples they are using. Rather than simply check a few cases and try to draw a conclusion like in naive empiricism, students at the crucial experiment stage now consider special cases and see if the conjecture 'breaks' when certain examples are used. These examples may include plugging in positive or negative numbers, numbers close to zero, or very large numbers to see what happens in cases far from the normal. However, students at this level still try to argue generalization from these cases.

At the generic example level, students begin to understand that an argument must be made separate from the examples that may have been used to gain intuition. The example itself serves as an entry point for the student through which some insight into the problem is gained. After gaining this insight, the student starts to make an argument in a more general fashion than relying on a small number of examples like in the previous two levels. Referring back to the quote from Aricha-Metzer and Zaslavsky (2019) from the introduction, a productive use of examples refers to students using examples at the generic example level of the framework described by Balacheff (1987).

I note here that I think examples are strictly positive experiences for students to use to develop intuition about proving. I position this article and this lesson for the inclusion of examples in student work. My goal is not to have students cease to use examples entirely, but rather to help students advance from the naive empiricism understanding of examples to the generic example level of example-use.

3. Context

This lesson is situated in an ITP course. Typically, students will have completed calculus 2 and will be introduced to proof concepts for the first time. This lesson is intended to be implemented early in the semester of the proof course to help develop students' intuition about examples and how to use them productively in their proving attempts. This lesson is built upon research findings from chapters 2 and 3 of the author's dissertation. The dissertation work is focused on how instructors of ITP courses perceive their students to use and understand examples in a proof context. Findings indicated instructors often recommended students who were struggling with the mathematical concepts involved in the proof refer to examples rather than jumping to a formal proof lacking examples of any kind. Further, instructors often recommended students focus on using examples more efficiently rather than exclude examples entirely. This research was used as a basis for the initial development of the interview protocol into a lesson for the enactment in an ITP course. As a note, the lesson described below was enacted before the final findings of chapters 2 and 3 of the author's dissertation. As such, some of the specifications of the lesson will be discussed in the final section for changes to be made.

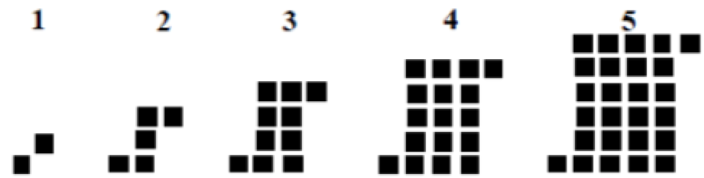
The materials used for this lesson are adapted from student work at a large public university in the southeastern United States. These student work samples are intended to be the heart of this task. Student work samples will be given to the class with the intention of learning from the work presented rather than assessing the work. The student work samples prompted thinking in the class and helped develop intuition for solving problems unfamiliar to the audience of the lesson. This type of task where student work is used to aid the learning process of a class is described by Bleiler and colleagues (2015) and Ko and colleagues (2016). Figure 16 shows the growing S pattern task to be given to the class.

Figure 16

Growing S Pattern Task

Growing S Pattern Task

Question 1. Consider the pattern below. How many square tiles would there be in the eighth step of this pattern?



Question 2. Write an expression for the total number of square tiles in the figure at an arbitrary step (n) of the pattern.

Question 3. Prove that your expression is a valid representation of the number of tiles at step n of the pattern. You may use drawings, words, numbers, and/or symbols for your proof.

Figure 17 describes the three student work samples adapted to the growing S pattern task. These student work samples were chosen to exhibit each of the three levels of example-use described by Balacheff (1987). I reiterate here that these student work samples are not meant to be assessment items for the audience of the lesson to grade. Instead, this task focuses on having students engage critically with the student work items and glean information about what parts of the provided solutions were helpful in the learning of the growing S pattern task.

Figure 17a

Nancy, Naive Empiricism

Nancy

Question 2. Write an expression for the total number of square tiles in the figure at an arbitrary step (n) of the pattern.

$$S = n^2 + 1 \in R, n > 0$$

S = number of squares, n = iteration

I noticed this pattern through the first several iterations of the task.

Question 3. Prove that your expression is a valid representation of the number of tiles at step n of the pattern. You may use drawings, words, numbers, and/or symbols for your proof.

Iteration	Squares	Correctness
1	$S = n^2 + 1 = 1^2 + 1 = 2$	Valid
2	$S = n^2 + 1 = 2^2 + 1 = 5$	Valid
3	$S = n^2 + 1 = 3^2 + 1 = 10$	Valid
4	$S = n^2 + 1 = 4^2 + 1 = 17$	Valid
5	$S = n^2 + 1 = 5^2 + 1 = 26$	Valid
8	$S = n^2 + 1 = 8^2 + 1 = 65$	Valid
10	$S = n^2 + 1 = 10^2 + 1 = 101$	Valid

The above expression is valid in each case above. This formula will hold for any value of n following this pattern.

Figure 17b

Carla, Crucial Experiment

Carla

Question 2. Write an expression for the total number of square tiles in the figure at an arbitrary step (n) of the pattern.

$$[(n + 1) * (n - 1)] + 2$$

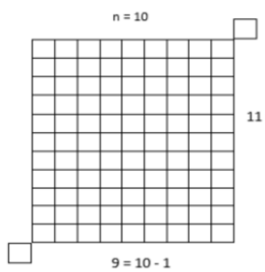
Question 3. Prove that your expression is a valid representation of the number of tiles at step n of the pattern. You may use drawings, words, numbers, and/or symbols for your proof.

n = 1

[(1+1)*(1-1)] +2 = 2

I tried the smallest case and a larger case. Since this formula holds for both, it makes sense for all cases.

n = 10



11 = 10 + 1

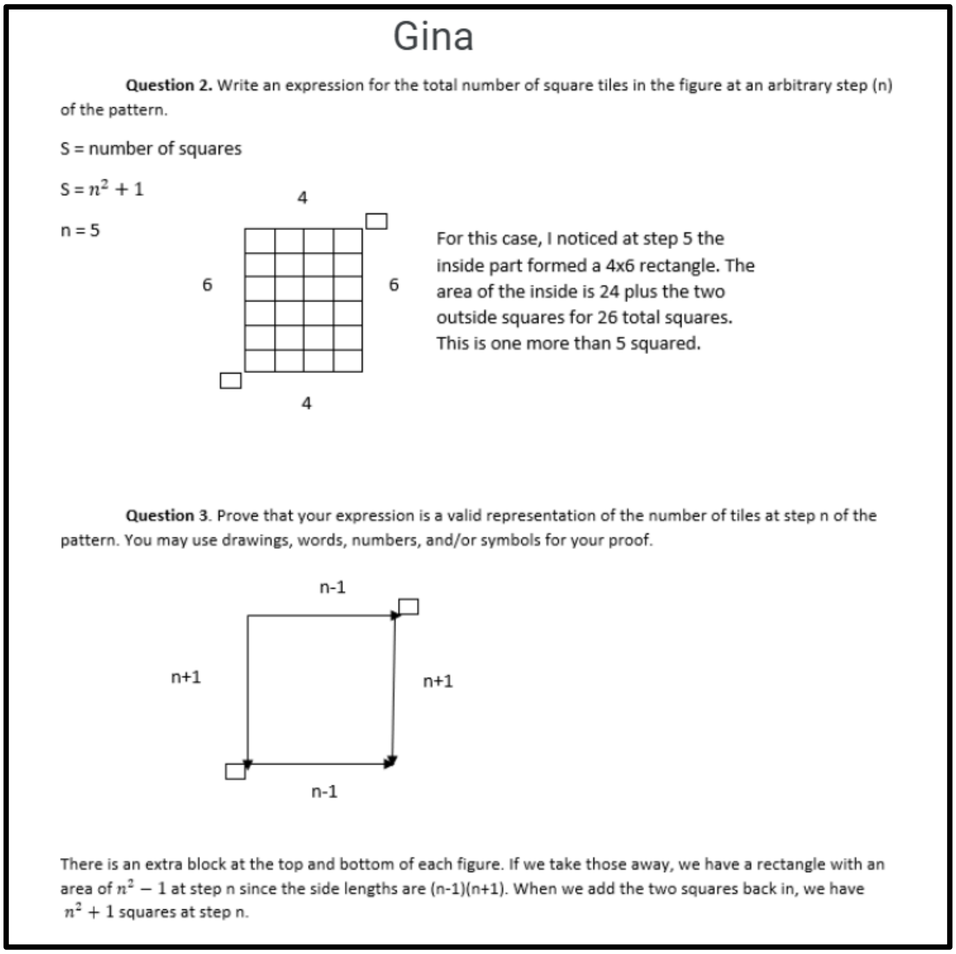
9 = 10 - 1

11 * 9 = 99 tiles

99 + 2 = 101 for the total tiles

Figure 17c

Gina, Generic Example



Gina

Question 2. Write an expression for the total number of square tiles in the figure at an arbitrary step (n) of the pattern.

S = number of squares

S = $n^2 + 1$

n = 5

For this case, I noticed at step 5 the inside part formed a 4x6 rectangle. The area of the inside is 24 plus the two outside squares for 26 total squares. This is one more than 5 squared.

Question 3. Prove that your expression is a valid representation of the number of tiles at step n of the pattern. You may use drawings, words, numbers, and/or symbols for your proof.

There is an extra block at the top and bottom of each figure. If we take those away, we have a rectangle with an area of $n^2 - 1$ at step n since the side lengths are (n-1)(n+1). When we add the two squares back in, we have $n^2 + 1$ squares at step n.

4. Justification

4.1 Findings from Research

This lesson is adapted from Kirby (2023). In this work, the student work samples above were given to instructors of ITP courses across the southeastern United States. Instructors were asked to respond to these student work samples and offer any insight or advice they may have. Although this problem was not seen before by 8 of the 11 participants of the study, they all expressed interest in this problem or something related as a good introduction to student thinking early in the semester. Some of the findings from Kirby's dissertation work include the tendency of instructors to guide their students from one level of the Balacheff (1987) example-use framework to the next rather than jump ahead to a formal proof with the example removed from the argument. Further, participants in the study consistently aligned their understanding of the student work and the level of sophistication they believed the students to be exhibiting with that of Balacheff's example-use framework. Finally, students who were often deemed struggling in their work were consistently encouraged to use examples to build intuition rather than provide a formal proof lacking examples.

This process was aligned with the recommendations from Desimone and Garet (2015) when discussing professional development. Desimone and Garet argued professional development is more productive and useful to practitioners when the teachers are met with their current practices rather than being thrown something new entirely. Kirby (2023) found common instructional practices of recommending students look for alternative proof methods when solving a problem. Further, practitioners of ITP courses often encouraged students to first consider an example to help build intuition for solving a problem. The current practices of practitioners of ITP courses is described in Kirby (2023) in more detail.

4.2 Designing a Lesson

There are two primary goals of this lesson. My first goal is for students. This lesson is intended to *encourage the use of examples when faced with a problem to help build intuition.* Aricha-Metzer and Zaslavsky (2019) argued examples used at the generic example level described by Balacheff (1987) were productive in helping the students successfully prove a mathematical claim. My second goal is for instructors. One of the findings from the author's dissertation work was that instructors of ITP courses were split on whether examples should be included in the formal written product provided by students. This lesson is intended to *show the usefulness of examples in helping students in an ITP course engage with proofs*. One of the expected outcomes of this lesson is the continued use of productive examples by students in the class, and the encouragement of example-use, when used productively, by instructors of an ITP course.

5. The lesson

This lesson is intended to be given within the first two weeks of an ITP class. This lesson was implemented with students and faculty of varying mathematical backgrounds at a small public institution in the southeastern United States. I will discuss how I implemented the task to an audience of students and faculty consisting of 29 individuals. I will primarily refer to the audience of this lesson as students even though several faculty members were present for the lesson. The faculty primarily observed the students in this lesson and were actively engaged without giving any answers or hints to the students. The layout of the classroom included desks in rows with a whiteboard at the front of the classroom. I had a PowerPoint prepared before the lesson with slides for each of the parts of the growing S pattern task. Further, I had slides prepared with each of the student work samples on them. Finally, I had slides at the end left purposefully blank to fill in with student ideas or thoughts after the conclusion of the lesson.

At the start of the lesson, I first gave question 1 of the growing S pattern task to all participants and had them do a think-pair-share activity for how to draw the shape. Once all involved had a chance to start thinking about the problem, questions 2 and 3 were shown to the audience. The audience was given about 5 minutes to quickly collect some thoughts about these questions in small groups of 2-3 before coming together as a class. Groups were assigned based solely on proximity. At this point, many of the students with fewer mathematics classes taken were typically struggling on how to approach the problem or even what it meant to "prove your expression is a valid representation…" The students with a more developed mathematical background (at least 2 proof classes) were beginning to have formulas developed. Further, every group with a student meeting these criteria started to look for some sort of inductive argument.

The class came back together at this point for a quick discussion. Some of the students raised issues to the whole class about what it means to prove something is valid, while other students said they had a formula they believed to be true but were unable to figure out how to inductively prove it. With these ideas stirring in their minds, I had the audience split into groups of 4-5 based primarily on proximity, but with at least one student with little mathematical experience and one student with a more advanced mathematical background in each group. After these groups were assigned, I gave each group one of the student work samples. Recall the three student work samples were Nancy (naive empiricism), Carla (crucial experiment), and Gina (generic example). In total, there were 6 groups, with each student work sample being given to 2 groups. The class was given the following prompt, "Your task is to read what this student has written on the handout provided to you and make sense of their work. Determine for your handout if you think the student successfully proved the problem. Why or why not?"

Students were then set free to work in their groups as they made sense of student thinking. As I walked around the class, I heard students in each group remarking about the use of examples. Whether it was Nancy's work or Gina's, students (particularly those who were struggling with the question) appreciated seeing some reasoning for how a formula was developed. Conversations continued in this fashion for the 10 minutes allotted for this section of the lesson as students worked in their groups to understand the work given to them and assess if the problem was proven by the student work sample. Students working with Nancy's work were the first to finish after about 5 minutes. Almost universally, all students in these groups agreed Nancy's work was insufficient for a proof. With the remaining time for this section, students working on Nancy's argument were asked to relate her work to theirs and see if they could offer Nancy any suggestions to improve her work. I made careful consideration to tell the students to not simply change Nancy's work to what they believed was correct, but instead have the students develop the work Nancy had presented to what they thought might meet their understanding of a proof.

After 10 minutes, I informed the students they were going to do a jigsaw activity. Each of the 6 groups sent 1-2 members to another group ensuring that the newly created groups had a student who originally worked on each of the student work samples. Students were now given the directions, "Teach your group members how your student approached this problem. List at least 2 productive things your student did, and 1 opportunity for improvement in the student work. Finally, mention whether you think this student proved the problem and your rationale for your answer." This part of the lesson would take most of the rest of the class period, leaving about 10 minutes at the end of the class for questions and discussion.

As students were working in their new jigsaw created groups, many found difficulty expressing productive things from Nancy's work. Similarly, several students found difficulty offering room for improvement in Gina's work. As the lesson progressed, most of the class agreed that all of Nancy, Carla, and Gina showed good intuition and problem-solving approaches by trying to start with concrete examples to deduce a formula. This helped reach my goal for students of the usefulness of examples to aid proof production. For both Nancy and Carla, students all agreed neither of these constituted a proof. All groups similarly had formality or a generalized argument as room for improvement in both Nancy and Carla's work.

Most of the groups agreed Gina's argument was sufficient for a proof. Students appreciated the example Gina started with and her generalized argument at the end. The more advanced students in the class offered a suggestion that Gina should look towards proof by induction for her room for improvement. This helped reach my second goal of having the more advanced students look for alternative proving techniques (proof by induction in this case). As the time was winding down on this part of the lesson, I asked each group their opinion of what a proof should include. This bled over into a full class discussion when I asked for guidelines for what makes a proof a proof.

The whole class discussion ended with students asking what proof should require and what I thought about the 3 student work samples. I left most of this largely open to the students, however I think this activity has a perfect lead in to a proof rubric activity described by Yee and colleagues (2018). Students were naturally curious about what a proof should or should not include, and many had different ideas. The class was split on whether Gina's argument would be improved or hindered if she took out her example to help answer question 2. Recall Gina had a generic argument where after plugging in the number 5 to her argument in question 2, she found the pattern of a rectangle to help answer question 3. The class was split on whether her answer to question 2 that focused on plugging in a specific number for her answer was sufficient for the problem. Some of the students argued that since this question did not specify for Gina to prove the claim her answer was sufficient. Others argued she should be consistent and use her answer from question 3 in a similar vein to how to answer question 2.

Further, many students (particularly the students with fewer mathematics classes taken), appreciated the work done by Carla and the examples she chose. Recall Carla had an argument at the crucial experiment level where she used two examples, one example that was her base case and another example which she thought was a larger example at the tenth step. Students appreciated the work of Carla and many started to make comments about how it was interesting to consider what would happen when n equals zero and they had not considered that. The advanced students also noted the similarity with the work Carla has shown with the idea of a base case in an induction proof.

6. Discussion

Productive use of examples can be a tool for students to use to help gain intuition and confidence when proving mathematical claims. Examples are used productively when students use the example to gain intuition about the generality for a deductive proof. Balacheff (1987) described this level of example-use as a generic example. I posit generic examples are strictly beneficial for students of all mathematical backgrounds to consider helping their proof production and understanding. The goal of this lesson was to help students see the benefits of examples and use them to aid their proof production and understanding. Additionally, this lesson was intended to help practitioners of ITP courses see benefit in the use of productive examples for their students. By building upon the research findings of the author's dissertation, this lesson was intended to meet instructors with practices like what they are currently doing, aligning with the recommendations for good professional development described by Desimone and Garet (2015).

The lesson presented in this article was implemented in a small public university in the southeastern United States. This lesson is meant to be implemented in the early semester of an ITP course to help students develop ideas for how to start proofs or what a proof even is. Students in this lesson made progress in understanding the role of examples and considering how to approach problems in multiple methods through evaluating student work samples. After the completion of the author's dissertation, a few alterations can be made to this lesson.

One of the primary findings in the author's dissertation was the split amongst faculty between whether examples should be included at all in a formal written argument. This notion was further mentioned by some select students when discussing whether Gina had successfully answered the problems in questions 2 and 3. This would present a great opportunity for a class discussion of the purpose of examples and whether examples can be included in proof. It would be beneficial to bring in practicing mathematicians and ask them if they use examples in their work to build intuition. Lockwood and colleagues (2016) interviewed mathematicians and found mathematicians commonly use examples in a productive manner to help build an understanding for a proof. This notion could help students see the benefit of starting with examples, even if they are not included later.

Another addition to the lesson would be to develop a sequence of lessons at the beginning of the semester for a better flow. This lesson assumes a base understanding for students to be familiar with at least the wording of a proof. A typical first class period in an ITP course should be sufficient for the background knowledge required for students to participate in this lesson. After this lesson, students should be thinking about the components of a proof and what a proof can look like. This would be a perfect lead in for a communal proof building activity like what is described by Yee and colleagues (2018). This activity would help students solidify what proofs should look like and what can be beneficial for the reading of other students' proofs.

Chapter 5

Conclusion

I set out to understand through this study how instructors of introduction to proof (ITP) courses perceive their students to use examples to aid their production and understanding of proof. Although much is known in the field of mathematics education research about how and why students use examples in proof (e.g., Aricha-Metzer & Zaslavsky, 2019; Ellis et al., 2019; Harel & Sowder, 2007), far less is known about how to help instructors of ITP courses get their students to use examples in a beneficial matter (Zaslavsky & Knuth, 2019). By investigating instructor perceptions of student example-use, I will help build the research base for how to aid instructors of ITP courses in getting their students to use examples productively.

An overarching framework surrounding my work consists of the example-use framework described by Balacheff (1987). This framework asserts three levels of example-use: naive empiricism, crucial experiment, and generic example. I interpret these levels of example-use to be along a continuum showing how students understand the purpose of examples in a proof context. Naive empiricism is apparent in student thinking when a student draws a conclusion based on a small number of examples and tries to make a general claim about a statement. The crucial experiment level of example-use moves further along the spectrum by instead being indicated in student reasoning when specific examples are chosen rather than seemingly random or irrelevant examples are chosen. Like naive empiricism, students exhibiting a crucial experiment level of understanding example-use attempt to make general claims about an argument from concrete examples. Finally, the generic example level of example-use is the most developed in the spectrum of student understanding of examples. This level of understanding involves students gaining intuition from an example to then make a generalized deductive argument when proving. Aricha-Metzer and Zaslavsky (2019) argued the generic example level of understanding is what constitutes a productive use of examples in students.

Chapter 1 - Introduction

In chapter 1 of this dissertation, I laid the basis for why this investigation is needed and defined terms used throughout the study. A common finding in the proof literature in mathematics education research is the difficulty students often have with learning how to prove or why a proof is even needed (e.g., Harel & Sowder, 2007; Healy & Hoyles, 2000; Stylianides, 2007). One issue contributing to this difficulty is the lack of understanding how or why to use examples to aid their proof tasks (e.g., Balacheff, 1987; Epp, 2003; Selden & Selden, 2003; Stylianides et al., 2017). This has led to a call for research from Zaslavsky and Knuth (2019) to investigate how to help instructors of ITP courses enhance their students' understanding and use of examples (p. 237).

My definition of an example comes from Zodik and Zaslavsky (2008), "A particular case of a larger class from which one can reason and generalize." (p. 165). More specifically, I care about how the student uses an example rather than the specific example the student used. Rather than investigating why some students plugged in the number 3 to help understand a problem, I am more interested in how the number 3 helped the student move forward in their mathematical understanding. Was the student successful in proving the problem? What about other students who also used the number 3? How can we as instructors help our students know to use the number 3 if it is helpful instead of the number 1 if it is not helpful? These questions drive much of my interest in example-use literature. Ultimately this led me to an overarching research question for this dissertation: How do instructors perceive the example-use of their students in an ITP course?

To answer this question, I conducted interviews with instructors of ITP courses across the southeastern United States. The transcripts from the interviews consisted of the primary data sources for this alternative style dissertation. Participants of this study were presented with 6 student work samples across 2 questions. Each of the 2 questions involved 3 students who had varied responses aligning with my interpretation of the Balacheff (1987) example-use framework. Participants were asked to respond to the student work samples in the context of a homework assignment being turned in to them. Responses included initial noticings, questions to ask the student, frequency of similar responses in their classrooms, how the participants might grade the response, or follow-up activities or suggestions to make to the student.

After responding to the student work samples, participants grouped the 6 student work samples however they deemed appropriate in as many ways as they could. These groupings were then asked to be named with justification for why each student was placed where they were. There were no restrictions on the number of groups required or how many students must be in a group.

For the final part of the interview, participants ranked the student responses from 1 (best) to 6 (worst) for what they understood the best attempt at a proof was. Participants talked aloud as they ranked the student work samples and were asked at the end of the interview how they would respond to the students they placed first and last respectively.

These responses were guided to include specific instructional moves or questions the instructor would do if these two students were students in their class. The subsequent three chapters of this dissertation analyze these interviews using qualitative analysis. Chapter 2 of this dissertation analyzed the broad perceptions of instructors to seek to help answer how instructors currently engage with examples used by their students. Chapter 3 takes a deeper dive into the responses of select participants of this interview and helps give clarity to the types of responses common amongst participants of this study. Chapter 4 takes the findings from chapters 2 and 3 and gives a starting point to help transition some research into practice geared towards current practitioners of ITP courses.

Chapter 2 - How Instructors Perceive Their Students to Use Examples for Proof

In chapter 2 of this study, I sought to answer the research question: How do instructors of ITP classes perceive students' understanding and use of examples? This work was primarily focused on the example-use framework described above by Balacheff (1987) through the lens of a social definition of proof adapted from the work of Stylianides and Stylianides (2017). The data from this study comes primarily from the responses of the participants to the student work samples, the grouping task, and the ranking task. This led me to four findings from this study.

The first finding of this study was the alignment with instructors' ranking with the levels of the Balacheff (1987) example-use framework. The students at the generic example level of understanding were largely ranked above all other students, while the students at the naive empiricism level were consistently ranked at the bottom by participants. This finding is promising for the intuitive alignment with how students use examples described by Balacheff. Of note, participants were not shown the framework described by Balacheff nor told the levels of example-use the student work aligned with. This intuitive alignment of practitioners with mathematics education research gives researchers a baseline to help develop instructional interventions and tasks appropriate for practitioners of ITP courses. Desimone and Garet (2015) argued a good practice for professional development is to meet practitioners where they are instead of coming from a disconnected research-heavy approach. This finding indicates instructors of ITP courses are aligned with the Balacheff example-use framework.

The second finding of this study was related to identifying the most common groupings done by participants in the grouping task. Participants frequently grouped student work by a perception of the students' understanding shown through the student work. Ten of the 11 participants grouped students using wording such as "needs improvement" for the students they perceived to not understand or not correctly answer the problem, and "critical thinking" for students they perceived to understand the problem. These groupings included 3 separate categories loosely aligning to the 3 levels of example-use described by Balacheff. This grouping again is promising for the intuition of instructors who were unfamiliar with the work described by Balacheff to align their ideas of correctness or validity with the levels of example-use. These groupings can be built upon to guide practitioners to the common ways students argue using examples that are already aligned with their intuition. Explicit instructional intervention targeting these levels of example-use can help practitioners guide their students to use examples more productively aligning with Aricha-Metzer and Zaslavsky (2019).

The second common grouping done by participants was whether they approved of the students' use of examples. There was an even split amongst participants who wanted to see an example in the final written product that was independent of length of time at the university, Carnegie classification of the university, public vs private university, and self-described teaching style. There was a slight difference between groups that is worth investigating further between participants with a mathematics Ph.D. and a mathematics education Ph.D. This led to my third and primary finding of the study.

My third finding in this study was a dive into the acceptance or rejection of example-use by students. By acceptance or rejection, I mean whether the participant expressed approval or disapproval for the inclusion of an example in the written work provided by the student. Participants in the disapproval category expressed concern with having students include their reasoning through examples in the written work provided in the interview. These comments were often met with perceptions of formality including, "There's really good ideas in all of them, but they need a little bit of massaging before they're, you know, ready for publication." The perception of what the purpose of an ITP course should be was evident through the comments made by participants in both groups. Alternatively, participants in the approval category of examples included comments like, "With beginners, I want you to be overly verbose and show why you know you're right… With learners, I worry more about whether you have every step down." I take this result, that instructors are split between acceptance and rejection of the inclusion of examples in formal written products, as my primary finding for this paper. If the field of mathematics education research wants to develop instructional interventions or materials for use in an ITP course, we as researchers should consider the beliefs and attitudes of practitioners in the development of materials.

My final finding in this study was how some of the participants responded to the student work. Irrespective of their approval of example-use, 9 of the 11 participants gave similar instructional strategies for students ranked 6th in the ranking task. These instructional strategies were to largely ask students to draw a related example to help build a sounder argument. The participants were split based on their approval of examples for how to respond to students ranked first. Participants in the approval category of examples often encouraged alternative proving methods to students. Participants in the disapproval category pushed for the students to construct a more formal argument not relying on an example. This finding further emphasizes the need to consider the perceptions of the practitioners when developing instructional tools or interventions to help students use examples more productively in an ITP course.

It is promising to see the alignment of instructors' intuitive ranking of students with the Balacheff (1987) example-use framework. This alignment will help guide future research in the field of example-use to develop instructional interventions and tasks appropriate for practitioners of ITP courses. Regardless of the acceptance of example-use in written products reviewed by participants, most participants expressed a desire to see examples written down and worked through to help struggling students. There is a disconnect between how students are encouraged to use examples when learning and the expectations for formal proof. I encourage future research to investigate whether this trend of approval or disapproval of examples is prevalent in other proof techniques such as induction or indirect proofs. Further, mathematics education researchers should consider the perceptions of practitioners when developing tools or interventions for use in an ITP course regarding examples.

Chapter 3 - Instructors Responses to Student Work Using Examples

In this study, I investigated the participants' responses further and answer the question: In what ways do instructors respond to student work samples involving examples in an ITP course? This study differs from chapter 2 by investigating how the instructors responded across the interview and dives deeper into the conversations I had with participants of the interview. The example-use framework described by Balacheff (1987) again frames most of the study. This study also incorporates the mathematical knowledge for teaching (MKT) framework and the related mathematical knowledge for teaching proof (MKT-P) framework developed by Ball and colleagues (2008) and Buchbinder and McCrone (2020) respectively for data analysis. Data in this study comes from participant responses across the whole interview, but primarily from the responses to student work samples and the ranking task. I will report on findings from this study across the three levels of Balacheff's framework: naive empiricism, crucial experiment, and generic example.

When participants were discussing student work at the naive empiricism level, an initial concern was raised for validity and precision in the writing. Most participants expressed concern for the validity of the argument with several calling out the students trying to do a 'proof by example' in the written work presented to them. Participants frequently responded with these student work samples with a strategy of asking the student to consider another example, often something like a monstrous counterexample as described by Stylianides and Stylianides (2009). Like my findings from chapter 2, it is promising to see the intuitive alignment with instructors' perceptions of their students' use of examples with Balacheff's (1987) example-use framework. This alignment can help serve as a starting point to develop instructional interventions or tasks building on what practitioners currently do in their practice. Interestingly, participants frequently expressed instructional strategies to help their students move forward to either a crucial experiment level of understanding or generic example level of understanding rather than a jump to a formal proof without examples entirely.

I had limited participant data on the crucial experiment level of student understanding due to the tendency of these students to be ranked in the middle for the ranking task. Most participants in the study found the work of Carla as promising and a "hint of n plus 1, n minus 1" as a push to generality of the statement. Like the naive empiricism level, common instructional patterns pushed participants to instruct the students to move to a generic example level of argumentation rather than a formal proof void of examples.

Finally, participants universally ranked a student at the generic example level of understanding first. Participants frequently expressed a desire to have these students consider alternative proof methods or focus now on formalizing the proof largely along the lines of example acceptance as described in chapter 2. These findings indicate that mathematics education researchers should consider developing instructional materials or interventions to guide students along the continuum of example use rather than a jump from using examples to a formal proof entirely to align with current instructional practices. This alignment with instructional practices matches a best practice for professional development described by Desimone and Garet (2015).

Chapter 4 - Engaging students in example-use in proof

In this chapter, I converted sections of my interview protocol into a lesson designed to be implemented early in the semester of an ITP course. This research to practice lesson was based on preliminary results from chapters 2 and 3 of this dissertation. I implemented this lesson in front of a group of 29 faculty and students at a small public university in the southeastern United States. This lesson was designed with the goal of encouraging example-use in students of ITP courses as well as help instructors see the benefits of productive example-use in proof production and understanding in their students.

This lesson is designed as a first attempt to answer the call from Zaslavsky and Knuth (2019) for instructional interventions or materials for ITP courses. Following the suggestions from chapters 2 and 3 of my study, this lesson builds on what instructors currently do when faced with students with varying understandings of examples. This research base aligns with how Desimone and Garet (2015) argued productive professional development can help practitioners. By giving the student work samples from my interview to students in the class I was teaching, the audience of my lesson was able to see multiple approaches to the growing S pattern task.

Collaboratively working through student work samples provided the class an opportunity to engage with thinking potentially foreign to them. Tasks utilizing student work to engage the class in discussions of mathematics are described by Bleiler and colleagues (2015) and Ko and colleagues (2016). Working through these student solutions aligned with what many of the instructors recommended Nancy and Aaron do in the consideration of alternative examples. Further, the class led to a discussion of the generality of a proof and the reasons most of the audience did not believe Nancy or Carla had a valid proof. This aligns with my finding from chapter 3 of participants frequently wanting their students to progress along the continuum of example-use one stage at a time rather than jump to a formal proof.

Finally, this lesson could fit in a potential lesson sequence at the beginning of the semester of an ITP course. At the conclusion of the lesson, the students in the class I presented to were having a discussion of what constitutes proof and how different proofs could look. This would be a great lead into a discussion surrounding what constitutes a proof through a proof rubric activity like what Yee and colleagues (2018) have described. Future lessons for instructional intervention should similarly investigate links between productive use of examples and instructional practices like a communal rubric building activity.

Comparison to existing literature

Much of the literature surrounding example-use in proof is focused on how and why students use examples when proving (e.g., Aricha-Metzer & Zaslavsky, 2019; Ellis et al., 2019; Zaslavsky & Knuth, 2019). Zaslavsky and Knuth (2019) argued more research is needed to help practitioners of ITP courses effectively guide their students to use examples more productively. Aricha-Metzer and Zaslavsky (2019) found that productive uses of examples were aligned with students using a generic example approach as described by Balacheff (1987). Buchbinder and McCrone (2020) furthered this finding and argued a goal of mathematics instruction should be to help students progress to deductive argumentation like what is found through generic examples.

The findings of this study partially align with how Buchbinder and McCrone (2020) argued instruction of higher-level mathematics courses should be handled. Some participants of this study valued and emphasized the role of examples and encouraged the use of generic examples by students to help gain intuition for solving problems. About half of the participants of this study instead took issue with examples showing up in the written work of students and instead insisted student work have a different sense of formality. The deductive argumentation described by Buchbinder and McCrone could be argued to align with either generic example-use or formalized arguments without the use of an example. I encourage future researchers to think further about how they view the role of examples in written work produced by students.

Lessons learned

I originally set out to conduct this study expecting the participants to appreciate the work of students at the generic example level, worry about students at the naive empiricism level, and have students in the crucial experiment level somewhere in between. My first takeaway of this study was that this expectation was mostly true. Practitioners were intuitively ranking student work in alignment with mathematics education research they were not shown nor familiar with. Some practitioners even gave definitions of the levels of student work aligning somewhat with the definitions described by Balacheff (1987).

My second lesson was something unexpected. I did not expect the split of participants that want to see examples in written work to any degree. This finding has produced multiple conference proceedings and discussions with scholars across the proof literature. The most eye-opening conversation I had was with another mathematics education researcher working on ITP classes who argued the purpose of an ITP course was to prepare students for formal, almost textbook style proofs. In writing these proofs, they argued students should be weaned off of using examples and shift to the style of writing common in publications and expected of other mathematics instructors. They posed an interesting question to me I am still considering, “When do you think (if at all) students should learn formal proof? Would it be harder to unlearn some of the suggestions you have in your presentation rather than start at the beginning with a formal proof?” I lean toward the inclusion of examples as a productive tool for students when writing proof, however I understand the opposing side. These comments help me consider some paths of future directions considering alternative perspectives.

Future directions and Recommendations for Future Work

If I could conduct this study again, I would have a few changes to make. Many participants took issue with the wording of Eric’s response and focused on his wording rather than the content of his mathematical argument. This detracted from collecting sufficient data for Eric and students at the crucial experiment stage. Further, I did not originally plan to focus on the data pertaining to instructor responses based on levels described by Balacheff (1987) so much of the data for students at the crucial experiment level of understanding is missing. Adding a direct question to have participants provide instructional strategies to each student would have been beneficial.

Future work in this field should consider how the acceptance of examples is tied to (if at all) the proof technique used. I would be very interested if similar trends to my findings for acceptance are found when investigating techniques such as mathematical induction or indirect proofs. It would also be interesting to investigate this group of participants’ views of the role of an ITP course. Many participants gave their views of the purpose of an ITP course unsolicited in my interviews. These views are likely one of the deciding factors of example-use acceptance. It may also be worth investigating how an instructors' view of the nature of mathematics relates to their perception of the use of examples in a proof setting. I recommend future research should investigate a similar sample to what I gathered in this study with some of the alterations to the interview protocol I discussed earlier.

Implications

The findings of this study give a baseline for mathematics education researchers to consider when developing instructional interventions or tasks for practitioners of ITP courses. The development of these instructional interventions should consider the possibility that many of the recipients or intended audience may be hesitant to include a push for example-use in their written work. Further, the audience of the instructional interventions are currently aligned with research including Balacheff's (1987) example-use framework allowing for a template to help design guidelines for productive use of examples. The lesson described in chapter 4 of this dissertation is one direction of how future work could help guide practitioners in aiding their students to use examples more productively.

Conclusion

This study sought to investigate instructor perceptions of student example-use and understanding in an introduction to proof course. Through interviews with instructors of ITP courses across the southeastern United States, I gained insight into answering this question by showing student work samples to the participants of my interview. Chapter 2 of this dissertation helped answer broadly how instructors of ITP courses felt about examples and how they group student's work using examples together. Participants of this study were split on whether to accept examples in the written work provided by students. Chapter 3 of this dissertation looked more deeply at the responses of participants to help give a look at how instructors would respond to students using examples in various fashions in their class. The results of this study indicated participants typically helped move their students along a continuum of example-use only moving away from examples when students reached the generic example level of example-use. Finally, chapter 4 of this dissertation helped make a connection between research and practice by building on the work of chapters 2 and 3 and designing a preliminary task to be implemented early in an ITP course. This task elicits student work samples to engage students of an ITP course in critical thinking of the usefulness of examples.

Proof is one of the hardest topics for many students to learn in mathematics. Mathematics education researchers have been studying ways to help students learn to prove more efficiently for decades (e.g., Balacheff, 1987; Healy & Hoyles, 2000; Stylianides, 2007; Zaslavsky & Knuth, 2019). This study helps provide a research basis for future work to design a research-based instructional intervention or tasks to be implemented in an ITP course. Future research should investigate how the findings of this study compare when looking across proof types including indirect proofs and mathematical induction. Further, additional research more closely targeted at instructional responses can develop the findings in this dissertation to better help develop instructional interventions.

REFERENCES

Alcock, L., & Inglis, M. (2008). Doctoral students' use of examples in evaluating and proving conjectures. *Educational Studies in Mathematics, 69*(2), 111–129. https://doi.org/10.1007/s10649-008-9149-x

Alcock, L., & Weber, K. (2010). Undergraduates' example use in proof construction: Purposes and effectiveness. *Investigations in Mathematics Learning, 3*(1), 1–22. https://doi.org/10.1080/24727466.2010.11790298

Aricha-Metzer, I., & Zaslavsky, O. (2019). The nature of students' productive and non-productive example-use for proving. *The Journal of Mathematical Behavior*, 53, 304–322. https://doi.org/10.1016/j.jmathb.2017.09.002

Balacheff, N. (1987). Processus de preuve et situations de validation. *Educational Studies in Mathematics, 18*(2), 147–176.

Balacheff, N. (1988). *A study of students' proving processes at the junior high school level*. Second UCSMP International Conference on Mathematics Education.

Balacheff, N. (2002). *The researcher epistemology: A deadlock from educational research on proof.* International Conference on Mathematics-Understanding Proving and Proving to Understand, February, 1–20. http://etnomatematica.org/articulos/Balacheff1.pdf

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59*(5), 389–407.

Bleiler, S. K., Ko, Y. W., Yee, S. P., & Boyle, J. D. (2015*). Communal development and evolution of a course rubric for proof writing*. In C. Suurtamm (Ed.), Assessment to enhance learning and teaching (pp. 97–108). Reston, VA: National Council of Teachers of Mathematics.

Boyle, J. D., Bleiler, S. K., Yee, S. P., & Ko, Y.-Y. (2015). Transforming perceptions of proof: A four-part instructional sequence. *Mathematics Teacher Educator, 4*(1), 32–70. https://doi.org/10.5951/mathteaceduc.4.1.0032

Buchbinder, O., & Mccrone, S. (2020). Preservice teachers learning to teach proof through classroom implementation: Successes and challenges. *Journal of Mathematical Behavior, 58*(February), 100779. https://doi.org/10.1016/j.jmathb.2020.100779

Buchbinder, O., Mccrone, S., Capozzoli, M., & Butler, R. (2022). Mathematical knowledge for teaching proof: Comparing secondary teachers, pre-service secondary teachers, and undergraduate majors. *International Journal of Research in Undergraduate Mathematics Education*. https://doi.org/10.1007/s40753-022-00187-8

Buchbinder, O., & Zaslavsky, O. (2019). Strengths and inconsistencies in students' understanding of the roles of examples in proving. *Journal of Mathematical Behavior, 53*(August 2018), 129–147. https://doi.org/10.1016/j.jmathb.2018.06.010

Buchbinder, O., & Zaslavsky, O. (2009). *A framework for understanding the status of examples in establishing the validity of mathematical statements*. Proceedings of the 33rd Conference of the International Group for the Psychology of Mathematics Education, 1(August), 225–232.

Cattaneo, K. H. (2017). Telling active learning pedagogies apart: From theory to practice. *Journal of New Approaches in Educational Research, 6*(2), 144-152. https://doi.org/10.7821/naer.2017.7.237

Chickering, A., & Gamson, Z. (1987). Seven principles for good practice in undergraduate education. *AAHE Bulletin, 39*(7), 3-7.

Cho, J., & Baek, W. (2019). Identifying factors affecting the quality of teaching in basic science education: Physics, biological sciences, mathematics, and chemistry. *Sustainability, 11*(3958), 1-18.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2007). *How and why do teacher credentials matter for student achievement?* CALDER Working Paper 2. Washington, DC: Urban Institute.

De Millo, R., Lipton, R. & Perlis, A. (1979). *Social processes and proofs of theorems and programs.* In T. Tymoczko (ed.), New Directions in the Philosophy of Mathematics, (pp. 267-285). Princeton University Press, Princeton, NJ.

Desimone, L. M., & Garet, M. S. (2015). Best practices in teachers' professional development in the United States. *Psychology, Society, and Education, 7*(3), 252-263

Ellis, A. B., Lockwood, E., Williams, C. C., Dogan, M. F., & Knuth, E. (2012). *Middle school students' example use in conjuecture exploration and justification.* 34th Annual Meeting of the North American Chapter of the International Group for the Psychology Of Mathematics Education, 135–142.

Ellis, A. B., Ozgur, Z., Vinsonhaler, R., Dogan, M. F., Carolan, T., Lockwood, E., Lynch, A., Sabouri, P., Knuth, E., & Zaslavsky, O. (2019). Student thinking with examples: The criteria-affordances-purposes-strategies framework. *Journal of Mathematical Behavior, 53*(January 2017), 263–283. https://doi.org/10.1016/j.jmathb.2017.06.003

Epp, S. S. (2003). The Role of Logic in Teaching Proof. *The American Mathematical Monthly, 110*(10), 886–899. https://doi.org/10.1080/00029890.2003.11920029

Ernest, P. (1998). *Social constructivism as a philosophy of mathematics*. State University of New York Press, Albany, NY.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Psychological and Cognitive Sciences, 111*(23), 8410-8415. https://doi.org/10.1073/pnas.1319030111

Harel, G. (2008). *What is mathematics? A pedagogical answer to a philosophical question.* In R.B. Gold & R. Simons (Eds.), Current issues in the philosophy of mathematics from the perspective of mathematicians. Washington, DC: Mathematical American Association.

Harel, G., & Sowder, L. (1998). *Students' proof schemes: Results from exploratory studies.* In A. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), Research in collegiate mathematics education III (pp. 234-283). Providence, RI: American Mathematical Society.

Harel, G., & Sowder, L. (2007). *Toward comprehensive perspectives on the learning and teaching of proof.* In F. K. Lester (Ed.), Second Handbook of Research on Mathematics Teaching and Learning (pp. 805–842). Information Age Publishing.

Harel, G., & Tall, D. (1991). The general, the abstract, and the generic in advanced mathematics. *For the Learning of Mathematics, 11*(1), 38–42. https://www.jstor.org/stable/40248005

Healy, L., & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for Research in Mathematics Education, 31*(4), 396–428. https://doi.org/10.2307/749651

Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction, 26*(4), 430–511. https://doi.org/10.1080/07370000802177235

Iannone, P., Inglis, M., Mejía-Ramos, J. P., Simpson, A., & Weber, K. (2011). Does generating examples aid proof production? *Educational Studies in Mathematics, 77*(1), 1–14. https://doi.org/10.1007/s10649-011-9299-0

Inglis, M., & Alcock, L. (2012). Expert and novice approaches to reading mathematical proofs. *Journal for Research in Mathematics Education, 43*(4), 358–390. https://doi.org/10.5951/jresematheduc.43.4.0358

Ko, Y., Yee, S. P., Bleiler-Baxter, S. K., & Boyle, J. D. (2016). Empowering students ' proof learning through communal engagement. *The Mathematics Teacher, 109*(8), 618–624.

Leron, U., & Zaslavsky, O. (2013). Generic proving: Reflections on scope and methods. *For the Learning of Mathematics, 33*(3), 24-30.

Lesseig, K. (2016). Investigating mathematical knowledge for teaching proof in professional development. *International Journal of Research in Education and Science, 2*(2), 253–270. https://doi.org/10.21890/ijres.13913

Lockwood, E., Ellis, A. B., & Lynch, A. G. (2016). Mathematicians' example-related activity when exploring and proving conjectures. *International Journal of Research in Undergraduate Mathematics Education, 2*(2), 165–196. https://doi.org/10.1007/s40753-016-0025-2

Lynch, A. G., & Lockwood, E. (2019). A comparison between mathematicians' and students' use of examples for conjecturing and proving. *Journal of Mathematical Behavior, 53*(February 2017), 323–338. https://doi.org/10.1016/j.jmathb.2017.07.004

Marrades, R., & Gutiérrez, Á. (2000). Proofs produced by secondary school students learning geometry in a dynamic computer environment. *Educational Studies in Mathematics, 44*(1–3), 87–125. https://doi.org/10.1023/A:1012785106627

Merriam, S.B. (2002). *Qualitative research in practice: Examples for discussion and analysis.* San Francisco, CA: Jossey-Bass.

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: Author.

National Governors Association Center/ Council of Chief State School Officers (2010). *Common Core State Standards for Mathematics*. Washington, DC: Council of Chief State School Officers.

Opdenakker, M., & Van Damme, J. (2006). Teacher characteristics and teaching styles as effectiveness enhancing factors of classroom practice. *Teaching and Teacher Education,* 22, 1-21.

Pascarella, E., Wolniak, G., Cruce, T., & Blaich, C. (2004). Do liberal arts colleges really foster good practices in undergraduate education? *Journal of College Student Development, 45*(1), 57-74.

RAND Mathematics Study Panel (2002). *Mathematical proficiency for all students: Toward a strategic research and development program in mathematics education* (DRU-

2773-OERI). Arlington, VA: RAND Education and Science and Technology Policy Institute.

Reid, D. A., & Knipping, C. (2010*). Proof in mathematics education: Research, learning, and teaching*. Rotterdam, the Netherlands: Sense.

Seifert, T. A., Pascarella, E. T., Goodman, K. M., Salisbury, M. H., & Blaich, C. F. (2010). Liberal arts colleges and good practices in undergraduate education: Additional evidence. *Journal of College Student Development, 51*(1), 1-22. https://doi.org/10.1353/csd.0.0113

Selden, A., & Selden, J. (2003). Validations of proofs considered as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education, 34*(1), 4–36. https://doi.org/10.2307/30034698

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

Sonnert, G., Sadler, P. M., Sadler, S. M., & Bressoud, D. M. (2014). The impact of instructor pedagogy on college calculus students' attitude toward mathematics. *International Journal of Mathematical Education in Science and Technology, 46*(3), 370-387. http://dx.doi.org/10.1080/0020739X.2014.979898

Steele, M. D., & Rogers, K. C. (2012). Relationships between mathematical knowledge for teaching and teaching practice: The case of proof. *Journal of Mathematics Teacher Education, 15*(2), 159-180. https://doi.org/10.1007/s10857-012-9204-5.

Stylianides, A. J. (2007a). Proof and proving in school mathematics. *Journal for Research in Mathematics Education, 38*(3), 289–321.

Stylianides, A. J. (2007b). The notion of proof in the context of elementary school mathematics. *Educational Studies in Mathematics*, 65, 1–20. https://doi.org/10.1007/s10649-006-9038-0

Stylianides, A. J. (2011). Towards a comprehensive knowledge package for teaching proof: A focus on the misconception that empirical arguments are proofs. *Pythagoras, 32*(1), 10.

Stylianides, G. J., Sandefur, J., & Watson, A. (2016). Conditions for proving by mathematical induction to be explanatory. *Journal of Mathematical Behavior*, 43, 20–34. https://doi.org/10.1016/j.jmathb.2016.04.002

Stylianides, G. J., & Stylianides, A. J. (2009). Facilitating the transition from empirical arguments to proof. *Journal for Research in Mathematics Education, 40*(3), 314–352. https://doi.org/10.5951/jresematheduc.40.3.0314

Stylianides, G. J., Stylianides, A. J., & Weber, K. (2017). *Research on the teaching and learning of proof: Taking stock and moving forward.* In J. Cai (Ed.). Compendium for research in mathematics education (pp. 237-266). Reston, VA: National Council of Teachers of Mathematics.

Weber, K. (2008). How mathematicians determine if an argument is a valid proof. *Journal for Research in Mathematics Education, 39*(4), 431–459.

Weber, K. (2010). Mathematics majors' perceptions of conviction, validity, and proof. *Mathematical Thinking and Learning, 12*(4), 306–336. https://doi.org/10.1080/10986065.2010.495468

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. https://doi.org/10.1080/00461520.2013.865527

Yee, S. P., Boyle, J. D., Ko, Y. Y. (Winnie), & Bleiler-Baxter, S. K. (2018). Effects of constructing, critiquing, and revising arguments within university classrooms. *Journal of Mathematical Behavior, 49*(November 2017), 145–162. https://doi.org/10.1016/j.jmathb.2017.11.009

Zaslavsky, O., & Knuth, E. (2019). The complex interplay between examples and proving: Where are we and where should we head? *Journal of Mathematical Behavior, 53*(October 2018), 242–244. https://doi.org/10.1016/j.jmathb.2018.10.001

Zaslavsky, O., Nickerson, S. D., Stylianides, A., Kidron, I., & Winicki-Landman, G. (2012). *The need for proof and proving: Mathematical and pedagogical perspectives*. In G. Hanna & M. de Villers (Eds.). Proof and proving in mathematics education (pp. 215-229). New York: Springer.

Zodik, I., & Zaslavsky, O. (2008). Characteristics of teachers' choice of examples in and for the mathematics classroom. *Educational Studies in Mathematics, 69*(2), 165–182. http://www.jstor.org/stable/40284541