The Rationality of College Mathematics Instructors: The Choice to Use Inquiry-Oriented Instruction

by

Mollee Shultz

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Educational Studies)
in the University of Michigan
2020

Doctoral Committee:

Professor Patricio G. Herbst, Chair
Professor Hyman Bass
Professor Kai S. Cortina
Professor Vilma Mesa

Mollee C. Shultz

mollee@umich.edu

ORCID iD: 0000-0001-7798-7390

Dedication

For Michael and Benjamin

Acknowledgements

There are so many people who made me the person and scholar who wanted to enter this journey. I thank my first teacher, my mom, for instilling a love for reading and writing as she homeschooled me. My dad, for always reminding me to "Learn your math and science!" and being my first mathematics teacher. Mr. Feuling, for being my first online math teacher and introducing me to Bohemian Rhapsody and Mountain Dew. Mr. Linn and Mr. Lauer, the 9th grade teachers who taught me lessons about reading, writing, speaking, philosophy, history, love, and life. Both also showed that these things could be taught with polar opposite teaching styles. My advisor Professor Balof at Whitman college, for sharing a joy for math, and Professor Gordon because he taught the vast majority of my math classes and made me understand what it meant to struggle with mathematics. Professor Laura Schueller, Linda Patton, and Emily Hamilton, thank you for existing as women in mathematics because having you as professors helped me believe that I belonged. Thank you to Trigger and Jingle for being the best sisters and friends who I can talk to anything about.

I have met and interacted with so many amazing people during this program that it is difficult to sum up in a few pages. To all my MRun friends, I look forward to few things as much as cross country and track seasons. Thanks for all the long car rides, Tuesday night workouts, and friendly competition. I wouldn't have made it through this program with my sanity without this club. To Elizabeth and Fit4Mom ladies, having this mom group has made all the difference. Whether it's in person or over Zoom, talking with you all leaves my whole heart in a fuller, happier state.

To all the math women at MSU - Reshma, Rani, Charlotte, Allison, Jessica, and Sara, thank you for your help with my research and drinking from baby bottles and spitting pacifiers with (against?) me. Rani, you're such a role model and I hope to do such a good job of straddling the math and math ed worlds in the future. I used your dissertation outline as a model for mine, and I cited you somewhere in here! To math guys at MSU – Leo, Harrison, Abhishek, Wenchuan, Sanjay, Hitesh, and Tyler, thank you for your help with my research and your friendship with Michael and me.

My amazing cohort has inspired me in all my classes and given me the foundation for my work. Darrell, thank you for all the great food, and being an example of someone who is quiet and yet has a strong presence. AC, the sweater you made for Ben is still and will probably always be the cutest article of clothing he's ever owned. I loved being able to work alongside you in an SFL journey. Laura-Ann, Ebony, Rosie, Matt, Ashley, Annie, Emily, Crystal, Naomi, and Darrius – you guys are forces of nature and I'm so impressed every time I see any work you've done. I can't wait to see what you continue to do. Anne, Colby, Lauren, Lindsey - I learned so much from you math ed people that were more experienced in the program. Emily, Amber, and Stacey, I love that we regularly were able to make time for lunches and hear about each other's very different routes through our programs. Amber, your support and example was so helpful while I was pregnant and first figuring out how to be a mom in academia. Xue Ying, Leo, and Stacey, it was fun having people to become parents with and go through all the stages around the same time.

I could not have produced this without the guidance of so many accomplished faculty. Mary, thank you for helping me publish my first paper and giving the most productive, helpful writing feedback. Vilma, thank you for including me in the IBL project in my first year. I learned how to use Word, how to write an interview, gave a first conference presentation, and had my first experience coauthoring a paper under your guidance. You have spent an incredible about of time giving detailed feedback for every major piece of work I have done during this program. Kai, for meeting with me and answering emails whenever I felt like I hit a dead end in this analysis. And Hy, for always being genuinely interested in ideas and engaging them in meaningful ways.

To the GRIP lab, working alongside you all has taught me how to do research. Amanda, I've been amazed from the first day at your productivity and ability to problem solve. Ander, I feel like you, Cassie, and Chloe laid down a blueprint for my family planning through the program. Also, I always thought the way you participated in GRIP meetings was the perfect amount – never monopolizing the conversation but adding extremely valuable tidbits when appropriate. Nic, I am in complete awe of your generosity with your time and care for your friends. I honestly cannot think of a more selfless person. There were so many times, during my first year and after, that I was completely confused by a theory, a project, or something stupid, and you happily talked through it all until I finally understood it. Throughout the program, you have listened to me vent about a huge variety of issues, and always have the perfect thing to say. I've watched many other newcomers lean on you in the same way. Claudine, you inspired the equity portion of my literature review with one small comment during a GRIP meeting. When you speak at meetings or otherwise, you don't have to be loud or aggressive. I've noticed that people have learned the value of your ideas, and want to hear what you have to say. Inah, you've been such a great friend and you're as selfless as Nic with your time. You have saved Ben's and my butts twice in the last couple months. Both Emanuele and Inah, there have been countless times when I've felt stuck and asked a simple question regarding statistics or using some

software, and you've taken the time to solve my problem. Emanuele, you continue to blow my mind, I'm often surprised (though I shouldn't be by now) by a hidden skill, or some hilarious comment I wasn't expecting. You and Lauren basically fueled the first few weeks of Ben's life. Mike and Emanuele, thank you for realizing my childhood dream of having real friends to play Age of Empires with. Mike, thanks for coming to Michigan and being officemates to really make our academic journey come full circle. I feel like we share a lot of perspectives. Sometimes I'll look over at you in a GRIP meeting when I feel like I can't express myself and know you get my perspective even if neither of us are any good at voicing it. Saba, (honorary GRIP member), I'm so grateful that I met you six years ago and we have come all this way together. You've clearly found your niche in the math education community and I couldn't be gladder that someone like you is doing the work you're doing.

To the leader of the GRIP team and my advisor, Pat, thank you for all you've done to foster my development the past few years. You've given me opportunities to put conference papers and presentations together, work on grants, and a variety of projects that fit my interests, and offered the feedback that was often challenging to hear but pushed products to be better. You're regularly thinking not only of how I can contribute to the lab, but also how different opportunities can contribute to my trajectory as a professional. Thanks for offering support beyond the work, from offering to equip my office for an infant to buying Ben blocks, Duplos, and the Frog and Toad stories. During the last few days of revising this dissertation and first few days of the COVID-19 chaos, your multiple offers of help were so appreciated. It means a lot that you care for your students and their lives outside the work place.

Perhaps the people to whom I owe the most thanks are the participants of this study. All were busy academics who took hours of their own time to help someone they had not met before.

I sent many emails to already teeming inboxes, and people were amazingly responsive. Reading through the open responses, people clearly put time and effort into giving thoughtful reactions. Finally, to Michael, thank you for your unconditional love and support. Thank you for not resenting and embracing that I can be a workaholic, for valuing our relationship and giving me the perspective that I need to not take all this too seriously, and for watching Ben countless hours so I could work and write. Or watching Ben so I could run or do other things that would help my mental health and enable me to work and write. Thank you for providing the motivation for me to apply to this program in the first place and editing all my applications with your sociology-math double major lens. This would not have been written without you, not only because of your emotional and logistical support, but because of the things you do and say as I've watched you teach in a math department the last five years. Some of the thoughts and biases I've developed, for better or for worse, are based on your experiences. Perhaps most significantly, there have been ways that I've learned to be in a relationship that transfer to the work done in this study. Lessons I've learned from communicating with you about how to really listen, how to be direct, how to see others' perspectives, have transferred to how I communicate with coworkers, respond to feedback, and write down my thoughts. You and Ben are everything to me and I'm so thankful to have you in my life.

Preface

My first experience learning calculus was in a community college. My professors were organized lecturers, patient explainers, and fair exam-writers, but I always felt like the rest of the class seemed to be in conversation with the professor while I couldn't keep up. I tried not to let it compound the insecurities I felt from being a 16-year-old girl. I spent a couple quarters feeling like an impostor, until I realized I actually performed better (by whatever standard was set by class averages on tests) than the peers I assumed were smarter than me. I realized that I learned by doing the homework and doing practice problems on my own or at the tutoring center. Typically, class time was only valuable to me after I struggled with the material on my own and brought back questions to class.

The more I have spoken with peers pursuing mathematics, the more I have felt like my experience is shared. While lectures are one way to ensure a professor "covers" the content of a course, and there are better and worse ways to lecture, I cannot believe that it is the most valuable use of the entire class time. This dissertation is a study of instructors' uses of other pedagogical strategies in college mathematics courses, and what is involved in their decisions to use them.

Table of Contents

List of Tables

List of Figures

List of Appendices

Abstract

This study of inquiry-oriented instruction (IOI) explores what inquiry-oriented practices are used by college mathematics instructors, and what relationships there are between their use of those practices, their beliefs about students' mathematics learning, and their recognition of professional obligations. I offer a conceptualization of inquiry-oriented instruction in which IOI practices documented in the literature are organized by the theory of the instructional triangle (Cohen, Raudenbush, & Ball, 2003), which pays particular attention to instruction as transactions of content between teacher and students. The INQUiry-Oriented Instructor REview (INQUIRE) instrument was developed on this conceptualization and used to gather data on the frequency that instructors report using inquiry-oriented practices. Professional obligations of mathematics teaching include the responsibilities that instructors have towards various stakeholders, including the institution, the individual student, mathematics as a discipline, and society (Herbst & Chazan, 2012) and instructors' recognition of these obligations was hypothesized as playing a role in explaining the use of IOI practices. A modified version of the PRofessional Obligations Scenario Evaluation (PROSE) instrument, a scenario-based assessment, was created for this study to be used with college mathematics instructors. In addition to developing the INQUIRE and PROSE instruments, this study incorporated an existing beliefs instrument (Clark et al., 2014) to measure instructors' beliefs on students' mathematics learning. I used factor analyses to confirm the hypothesized inquiry-oriented practices in the instructional triangle framework and hierarchical cluster modeling to reveal patterns of inquiry-oriented practices among instructors. I found that instructors reported using seven distinct sets of practices, and instructors grouped into four

different clusters based on their pattern of use of these practices – revealing different characterizations of IOI. This finding has implications for future research of IOI, showing that characterizing IOI as a singular pedagogy is problematic; rather, there are different types of IOI that are grounded in content-specific interactions. The first cluster includes participants that report the highest use of teacher-student and student-student interactions, but not the highest use of the student-content practices of giving students opportunities to construct and critique claims or write proofs. The second cluster includes participants that report the highest use of the aforementioned student-content practices, and second highest use of all five other inquiry-oriented practices. The third cluster included participants that reported low use of all inquiry-oriented practices except the teacher-content ones of interactive lecture and hinting without telling, which they use at levels comparable to other clusters. These three clusters or characterizations of IOI are all juxtaposed against the fourth cluster, which had the lowest reported use of all seven practices. I used structural equation modeling to explore the hypothesized relationships. Past studies have reported inconsistencies between beliefs and practice; instructors' degree of recognition of the professional obligations helped explain why instructors may not always actualize their beliefs in the classroom. I found that learner-focused beliefs often predict the use of inquiry-oriented practices, but recognition of the disciplinary and interpersonal obligations can work in direct opposition of those beliefs – helping to explain why instructors sometimes do not instruct with IOI even if they believe it would be beneficial. These findings have practical implications for those wishing to shift trends in college mathematics instruction. Future work could use the INQUIRE instrument to link inquiry-oriented practices to student experiences.

Keywords: Undergraduate mathematics, inquiry-oriented instruction, inquiry-based learning, professional obligations, beliefs, INQUiry-oriented Instructional REview (INQUIRE), PRofessional Obligations Scenario Evaluation (PROSE)

Chapter 1 **Introduction**

Tertiary mathematics departments are faced with a growing demand to develop innovative instructional practices to address the needs of increasingly diverse student bodies and declining numbers of mathematics majors (Holton, 2001; U.S. Department of Education, 2006). Even when instructors give a high-quality lecture, students often do not grasp the main ideas the instructor intends to convey (Goodstein & Neugbauer, 1995; Leron & Dubinsky, 1995; Lew, Fukawa-Connelly, Mejía-Ramos, & Weber, 2016). That is, simply presenting the content clearly does not ensure that students will learn it, and evidence shows they often do not. Rather, some amount of activity on the part of the student beyond rote notetaking is beneficial. Students learn better from active, student-centered instruction in college mathematics (Kwon, Rasmussen, & Allen, 2005; Rasmussen, Kwon, Allen, Marrongelle, & Burtch, 2006). It is even better if the student activity takes more thinking on the part of the student that is beyond copying a procedure that has already been demonstrated. Stigler, Gallimore, and Hiebert (2000) found that, in U.S. classrooms, mathematics instructors tended to instruct by teaching routine procedures and letting students practice them, as opposed to other countries, such as in Japan, where students spent more time inventing, analyzing, and proving.

Researchers have argued that instruction that emphasizes finding a singular solution given an algorithmic procedure over reasoning and making connections between mathematical ideas limit students' opportunities to engage meaningfully with the mathematics (Boaler, 1998; Hiebert et al., 1997; Kilpatrick, Swafford, & Findell, 2001; Lesh & Harel, 2003). Moreover, meaningful participation in inquiry-based instruction has been linked to higher achievement and persistence for women and students of color (e.g., Boaler, 1997; Laursen, Hassi, Kogan & Weston, 2014). Mason (2001) reported this shift in the mathematical teaching practices at the tertiary level ICMI report advocating for "a revised didactic contract, seeking to balance developing competency with enculturation into mathematical thinking, rather than succumbing to student desire to minimize effort and simply be trained in requisite behavior" (p. 72). Researchers have reported that mathematicians have shown enthusiasm for teaching with inquiry-oriented methods (e.g., Johnson, Caughman, Fredericks, & Gibson, 2013; Wagner, Speer, & Rosa, 2007), and many universities are funded to support instructors in implementing IBL (e.g., the Educational Advancement Foundation funds the Academy of Inquiry-Based Learning).

Though the education literature has long advocated for more student engagement and inquiry, the majority of science, technology, engineering and mathematics (STEM) undergraduate courses continue to be taught in a lecture format where students primarily sit and listen (Stains et al., 2018). Advanced mathematics courses have been documented as frequently taught in a "definition-theorem-proof" lecture format (Weber, 2004). That is, instructors write some definitions on the board, relevant to whatever content is being taught. Then they write the theorem that uses the defined terms. Finally, they prove the theorem. In lower-division courses, this is often accompanied by examples of problems and solutions that rely on that theorem. Researchers have documented how complex and challenging it is to change instructional practice with educational reform movements (e.g., Andersen, 2011; Cohen, 1990), and the literature tends to reflect some amount of exasperation that instructors continue to ignore such recommendations. Holton (2001) suggested that university mathematicians tend to perpetuate the same teaching practices that they experienced as students. Researchers have found that even if instructors believe a given practice is best, they do not always enact it (DeFranco & McGivney-Burelle, 2001; Raymond, 1997; Skott, 2001). From my experience teaching and from observing my own instructors, I anecdotally observed that instructors who chose to continue teaching without innovative, active learning practices did so for good reason, not out of laziness or inability to change. For example, my topology instructor gave clear, well-prepared lectures with beautiful diagrams and board work. In a later semester, she used more active learning practices like having us give student presentations, so I knew she was capable and willing to use them. I think she chose to use lectures because she felt displaying the introductory content was valuable for multiple different reasons. From my student perspective, her drawings and explanations decoded our dense, convoluted textbook. To capture the complex reasoning behind instructors' practices, this study draws from the literature on instructor decision-making to understand why instructors use or do not use practices associated with IOI.

The investigation of two related issues can be distilled into two research questions: First, what exactly is inquiry-oriented instruction, and second, how do various resources, both individual and social, contribute to decisions to use or not use inquiry-oriented practices? Instructors may choose to use inquiry-oriented practices based on individual characteristics or dispositions, such as who they are, how much experience they have teaching, whether or not they believe the inquiry-oriented practices were effective, or based on personal experiences they had in the past. But there may be other social factors, such as the expectations from students or colleagues at their institutions, that also play a role in instructors' decision-making. Studies have not empirically investigated these factors, especially not together in one study. Before reviewing the literature on what has been found so far concerning individual and social resources, I clarify a central problem with the research on IOI which led to the first research question.

What is IOI

Amid all the recommendations to change, there are a wide range of definitions of what instruction with inquiry entails. Part of the confusion is that there are multiple independent tracks of theory, practice, and research that have led to inquiry-oriented instruction as a movement. Independent of the development of the Moore Method and the research on inquiry-oriented instruction and inquiry-based learning, scholars have stressed the importance of inquiry and discovery for students. Bruner (1960) wrote that an important aspect of teaching is fostering "excitement about discovery – discovery of regularities of previously unrecognized relations and similarities between ideas, with a resulting sense of self-confidence in one's abilities" (p. 20). Bruner reported that people who work on mathematics curricula contended that it was possible to design sequences of problems that allowed students to discover these relations and structures themselves. He praised the Committee on School Mathematics and the Arithmetic Project of the University of Illinois for designing methods that would allow students to discover generalizations behind particular mathematical operations.

Like Bruner, other scholars have argued that students that discovered knowledge would both understand it better and be more engaged in the mathematical process than students who were merely told what to think. Scheffler (1965) distinguished between a strong and a weak sense of knowing. The weak sense of knowing only entails believing something was true, while the strong sense means also being able to justify the truth of a belief. He suggested that a way for teachers to facilitate such strong knowing would be to use teaching methods that require discovery and problem-solving. Freire (1990) advocated for *problem-posing education* where "the teacher is no longer merely the-one-who-teaches, but one who is himself taught in dialogue with the students, who in turn while being taught also teach" (p. 67). He emphasized the importance of engaging students in "inquiry and creative transformation" (p. 71).

There are two other main tracks of this movement – one originating in the practice of college professors, and another from K-12 mathematics education researchers (Laursen & Rasmussen, 2019). R. L. Moore, the mathematician who inspired the Moore Method and much of the inquiry-based learning (IBL) instruction in U.S. college classrooms (Laursen, Hassi, Kogan, & Weston, 2014), took students on a "journey of discovery" where they were not allowed any textbooks, previous exposure to subject matter, or library resources (Parker, 2005). The IBL movement that has been built upon Moore's legacy has struggled to brand their style of teaching with a well-defined framework. Many practitioners prefer a less rigid version of the Moore Method. The Modified Moore Method (MMM), for example, includes many of the same elements as Moore's original courses but in a more relaxed way. It is based on providing students with definitions and having them work on exercises and proofs without common resources like textbooks and lectures, but with more collaboration and scaffolding. Chalice (1995), a mathematician who practices the MMM, advises sending students up to the board in groups of three instead of one at a time, giving explicit instruction on techniques on theorem proving, and giving exercises on definitions as a way to warm up and support understanding. The literature on IBL includes a variety of ambiguous definitions. For example, Laursen and colleagues (2014) defined inquiry-based learning as a method which "invites students to work out ill-structured but meaningful problems […] construct, analyze, and critique arguments, […] present and discuss solutions alone at the board or in small-group work, while instructors guide and monitor this process" (p. 407). Yoshinobu (2012) wrote about different levels of IBL, ranging from "active lecture" to full IBL, attributing value to any practices that increase student engagement.

In contrast, another track of the inquiry movement has been significantly influenced by the research done in K-12 classrooms. The definitions of IOI that have emerged from the research track of work are just as varied as the practice-based work. National Research Council (1996) characterized inquiry as including identification of assumptions, consideration of alternative explanations, and critical and logical thinking as a research-based recommendation for K-12 classrooms. In a review of the science education K-12 research on IBL, Pedaste and colleagues (2015) defined IBL as "an educational strategy in which students follow methods and practices similar to those of professional scientists in order to construct knowledge. (…) Inquiry-based learning emphasizes active participation and learner's responsibility for discovering knowledge that is new to the learner" (p. 48). Their idea of IBL involves having students follow practices similar to professionals in the field is similar to what the theorists said, though what exactly those practices entail is not necessarily common knowledge.

Also, from the K-12 arena of research, the work done by Paul Cobb, Erna Yackel and colleagues was highly influential to the work being done on IOI in mathematics education (Laursen & Rasmussen, 2019). They highlighted mathematics instruction that incorporated constructivist pedagogies and activities that developed students' thinking strategies (Cobb et al., 1991) and studied how group work centered around problem solving could give rise to opportunities for collaborative dialogue, and, in turn, learning opportunities (Yackel. Cobb & Wood, 1991). They pioneered the ideas that led to two key pillars for social interactions in IOI classrooms, that "(1) students share their thinking, and (2) students orient to and engage in others' thinking" (Laursen & Rasmussen, p. 133). Laursen and Rasmussen (2019) proposed two more pillars of IOI based on more recent research, including "(3) helping students deepen their thinking, and (4) building on and extending student ideas" (p. 134). They state that, "the role of an IO instructor is multi-faceted with practices and routines that go well beyond those required for lecturing and the dissemination of knowledge" (p. 134). However, what those practices and routines consist of and whether all four pillars can be accomplished simultaneously is yet to be systematically investigated.

Laursen and Rasmussen (2019) posited that, though IBL (stemming from the Moore Method) and IOI (stemming from the K-12 research) have different historic and research foundations, they both emphasize similar instructional practices – which are what this study attempts to both discover and measure. I use the term *inquiry-oriented instruction* (IOI) in this dissertation to describe any or all instructional practices documented in the IBL and IOI literature aimed to increase the students' engagement, with the goal of discovering what inquiry-oriented practices are reportedly being used and why by college mathematics instructors. For this work, I choose to use the term IOI over IBL due to the focus on instruction rather than on student learning. Investigating the composition and use of IOI is a first step to understanding what opportunities IOI affords. This line of investigation is the type of work that the field of mathematics education can emulate with other forms of relevant instructional strategies. This study provides a blueprint of how to bring theories and practices from other levels of educational research and apply it to the college mathematics context.

Bridging Three Areas of Research

The study of what IOI is and what factors predict its use sits at the junction of three lines of research: work in Research in Undergraduate Mathematics Education (RUME), research on inquiry-based learning, and research on teacher decision-making. It contributes to the RUME community by bringing a practical, relevant topic (IOI) in communication with theories not as well-known in the college-level literature. Many of the studies on the use of instructional strategies done at the college level tend to report what is happening, without a clear theoretical framework to organize what or why things are happening (e.g., Rasmussen et al., 2019; Stains et al., 2018). This information can be useful to stakeholders like mathematics departments in a technical manner, but such literature lacks a blueprint that future researchers can use to build off the previous knowledge. Cai and colleagues (2019) wrote of the importance of the theoretical framework in guiding any study's hypotheses and methods, and its role in guiding the claims that can be made from the findings. They explain that the development of the theory is what advances the field's knowledge.

The theoretical frameworks on teaching used in this study and some of the IOI literature used to develop the study's instruments come from extensive work done at the secondary level. This study translates and applies it to the college level. This study also posits a metaframework for organizing the constructs involved in IOI. Both these efforts demonstrate how this study not only contributes practical ready-to-use results, but also leverages the knowledge generated by mathematics educators that can be used in future research specifically in the RUME community. The theoretical frameworks on teaching can also contribute to the literature on teacher decision-making.

The literature on teacher decision-making in mathematics classrooms has largely stayed in the realm of K-12 education and used frameworks that do not consider social resources (Stahnke, Schueler, Roesken-Winter, 2016). In a review of 60 articles on teacher decision-making in mathematics education, Stahnke and colleagues (2016) identified only four that were conducted in higher education contexts. In college, instructors' decisions are contingent on different variables and instructors have claimed to have more autonomy in their classrooms than at K-12 settings. Thus, there is reason to believe that studies in college settings could reveal a different set of findings than in other settings. The main findings from Stanke and colleagues' (2016) review concerned teachers' knowledge, skills, beliefs, and dispositions, but nothing along the lines of norms or obligations inherent to their professional contexts. This study brings in additional frameworks to explain teacher decision-making, as well as providing results from a population that has been studied very little. The decision-making literature provides a frame that has not been used to explain research on teaching with innovative instructional practices.

The research on inquiry-based learning or inquiry-oriented instruction could benefit from more study of the role of the instructor in the college mathematics context. As explained in the previous section, there are many practitioner accounts and recommendations for teaching with inquiry in the college context, but there have not been studies of what is occurring broadly in classrooms in the college context. This study provides a framework for what the components of IOI entail in the college context that could also be transferrable and comparable in K-12 settings. I suspect some of the practices that are more innovative and infrequent in college classrooms (e.g., interactions between students, more scaffolding of activities) are expected more regularly by students in K-12 settings. Thus, the overall understanding of what inquiry-oriented practices are occurring and why they are occurring at any level, can benefit from the theoretical and methodological work done in this dissertation. I turn to the outline of how I organized my efforts to build bridges between the three areas of RUME, teacher decision-making, and research on inquiry-oriented instruction in this manuscript.

Dissertation Overview

In this first chapter, I introduced the multiple threads of work behind the term IOI and the lack of clarity on what IOI is. The second chapter outlines existing literature on inquiry-based instructional practices, including the potential opportunities for equity-related implications that this work would enable researchers to investigate, and an organization of the literature based on the instructional triangle. It ends by reviewing literature on instructor beliefs that may have relevance to the use of IOI. The inconsistencies found between instructors' beliefs and practice motivates the framework outlined in the third chapter. The framework of practical rationality gives a lens to understand the decision-making of instructors to use components of IOI or not. I outline how each of the four professional obligations might play a relevant role, and end with research questions and a hypothesized model for interactions between IOI, beliefs, and professional obligations. In the fourth chapter I explain the motivation and design behind the instruments used in this study, along with the plans for sampling and analysis. In the fifth chapter, each section focuses on the scale development of a separate instrument. I use these scales to determine the relations between constructs in the sixth chapter, where I report results on what IOI practices are used and, from the cluster analysis, what patterns of IOI exist. Given these results, I present how instructors' beliefs and recognition of professional obligations predict the practices and patterns of practices. In the final chapter, I discuss the theoretical, practical, and methodological implications of the results. I describe changes I would like to make in future iterations of the INQUIRE instrument, as well as general future directions for research. The chapter ends with a discussion of limitations of the study and a final conclusion.

Chapter 2 **Review of the Literature**

This section first reviews the literature on IOI and suggests a way to organize the diversity of actions often included among inquiry-oriented practices. One reason why the organization of inquiry-oriented practices is essential for the development of research on IOI is because of the equity-related implications of using IOI. The field cannot yet parse apart why studies sometimes yield conflicting results. I then turn to the literature on teacher beliefs, as beliefs have traditionally been used to help explain differences in classroom practice. The discussion of beliefs motivates the theoretical framework that follows in the next chapter, where I argue that though beliefs are a significant individual resource, they do not account for how individuals relate to the environment in which instruction deploys. The framework helps provide a more nuanced and balanced explanation for instructors' decisions.

Opportunities in Inquiry

I focus on college-level mathematics because there is attrition of students interested in mathematics during undergraduate studies (Matyas & Dix, 1992). Using the metaphor of the mathematics pipeline, the pipes become particularly leaky during college and it is said that inquiry-oriented instructional practices may help stop this leak. In the pipeline metaphor, each separate pipe represents a level of schooling, such as elementary school, secondary school, or graduate school. When the pipes don't fit snugly, the liquid running through will leak out – more and more at each junction. The liquid is a metaphor for the students that begin to abandon studying a certain subject as they advance in their studies, potentially due to a reduction in opportunities from instructors to successfully identify the student as a doer of mathematics. The

Committee on Women in Science and Engineering of the National Research Council reported that there are increasing numbers of U.S. students who initially show interest in STEM fields but leave those fields while pursuing their undergraduate degrees (Matyas & Dix, 1992). For example, though there are similar amounts of girls and boys enrolled in mathematics in high school (National Center for Education Statistics, 1997), rates of attrition increase for girls as they progress in higher education (Herzig, 2004). Similar patterns of attrition exist for students of color pursuing mathematics (Herzig, 2004). Research on IBL has shown potential to decrease these achievement gaps. For example, Laursen et al. (2014) found that use of IBL eliminates the gender gap that exists in lecture-based courses. In both IBL and non-IBL courses, women's grades were as good as their peers, but the self-reported learning and affective gains of women were lower in non-IBL classes. The women in IBL classes generally reported higher confidence and personal interest in pursuing mathematics in the future. Their study suggested that IBL approaches provided students with learning experiences that were equally valuable for women and men.

A potential source for inequities in doing mathematics comes from tendencies to ascribe mathematical authority based on status rather than reasoning. Yackel and Cobb (1996) illustrated how one student attempted to solve a mathematical dispute by starting a conversation about who was the smartest and who had the best pencil. Even professional mathematicians sometimes rely on other indicators (such as the status of other mathematicians who reviewed the proof) than mathematical content to judge the soundness of published proofs (Weber & Mejía-Ramos, 2011). Thus, emphasizing practices that encourage students to engage with each other and the content to establish truth rather than relying on arbitrary authorities such as teachers and the textbook can potentially lead to a more equitable learning environment.

Disaggregating student data has shown that IOI helps to narrow some achievement gaps. Eddy and Kogan (2014) conducted an experiment to evaluate the impact of an active learning intervention in a college introduction to biology course. The course was taught for three terms in a traditional lecture format and three terms in which more active learning was included. During the active learning weeks, the initial coverage of content was offloaded into preparatory homework, while class time was used for activities in informal small groups that reinforced main ideas and allowed students to practice critical thinking skills. Students practiced answering questions similar to what would appear on exams using classroom-response software. The researchers found that the active learning was beneficial to students, and more notably, the intervention disproportionately helped the exam performance of black and first-generation college students compared to the rest of their peers.

Although general inquiry-oriented practices have shown to increase equity in terms of gaps in exam performance, not many studies have parsed out components of inquiry-oriented practice for quantifiable study. Different practices have shown to have different impacts on some student subpopulations. For example, Kim (2002, 2008) determined that Asian-American students did not benefit from talking about problems out loud, as opposed to being allowed time to think silently, the same way that their European-American student counterparts did. Some studies show that inquiry-oriented practices do not create equitable outcomes for performance, without any clear explanation why (e.g., Johnson, Andrews-Larson, Keene, Melhuish, Keller, & Fortune, 2018). By addressing the issue of what IOI is and gathering data on what practices instructors are using, I aim to provide methods that can further research in issues of equity. Inconsistent studies on the relationship between IOI and equity are part of a larger group of studies of IOI and outcomes. They all share the central problem that the inquiry-oriented instruction is not well-defined, so it is not clear how results can be generalized or what is causing the conflicting outcomes.

Studies of Inquiry-Oriented Practices

Much of the literature on inquiry-oriented practices in mathematics classrooms has been focused on IOI as established, wholistic pedagogy. Often studies focus on the design and implementation of a particular task or inquiry-based curriculum in a given context (e.g., Kwon, Rasmussen, & Allen, 2005; Larsen, Johnson, & Bartlo, 2013; Lockwood, Johnson, & Larsen, 2013; Rasmussen & Kwon, 2007; Stephan & Rasmussen, 2002; Wawro, Rasmussen, Zandieh, Sweeney, & Larson, 2012). When the research is not focused on design, its goal is often to investigate student outcomes from IOI courses, such as ability to think flexibly and innovatively (e.g., Boaler, 1998), grades in subsequent courses (Kogan & Laursen, 2014), ability to construct proofs (Smith, 2005), and student retention (Laursen et al., 2014). In all these studies, what counts as IOI was not analyzed. As such, it is not only difficult to know how the results can be replicated or generalized, but also to support practitioners who are left without a clear roadmap of how to establish similar outcomes. The lack of clarity around what IOI might have contributed to the lack of large quantitative studies on the subject.

The methods used to study inquiry-oriented practices have tended to be qualitative unless they are student-focused, in which case they may also use quantitative methods. Some studies use interviews with instructors (e.g., Johnson et al., 2013; Mesa, Shultz, & Jackson, 2019) and some also use classroom observations (e.g., Wagner et al., 2007). The studies have often been case studies of one or a handful of instructors who profess to use inquiry-oriented instruction (e.g., Boaler, 1997, 1998; Johnson et al., 2013; Wagner, Speer, & Rosa, 2007; Wawro et al., 2012). The few studies that are quantitative often use data that do not clearly articulate the actual instruction occurring. The studies are often more focused on student outcomes from the instruction (e.g., Kogan & Laursen, 2014; Laursen et al., 2014) or they characterize IOI in limited ways. For example, Marshall and Horton (2011) report on a quantitative study using 110 mathematics and science classroom observations and comparison tests to explore the relationship between inquiry-based instruction and student higher-order thinking. They measured inquiry-based instruction based on how much the teacher waited to give explanations until students had a chance to explore a concept and whether the students added to the explanation (Marshall & Horton, 2011). While this could be one component of IOI, it is a limited criterion for measuring IOI in its entirety.

In another example of a quantitative study with a limited characterization of IOI, Wilkins (2008) used self-report surveys and structural equation modeling to explore the relationship among teachers' knowledge, attitudes, beliefs, and practices. His study included 281 elementary teachers. He conceptualized IOI as a unidimensional construct, including items like "investigate mathematics through children's literature" and "play mathematics games" (p. 159), which would likely not be relevant to IOI at the college level. He did have some components that could apply to other levels of schooling, such as, "Work on problems for which there is no immediately obvious method of solution," and, "design or implement their own investigation" (p. 159), but additional work would have to be done to turn those items into a scale. Some studies not only treat IOI as a unidimensional construct, but also only use a couple ill-defined items to measure it. Engeln, Euler, and Maass (2013) conducted a study of the inquiry-based learning beliefs and practices of 917 secondary science and mathematics teachers across 12 European countries. The study measured instructors' use of IBL based on self-report responses to the two items, "I regularly do projects with my students using IBL," and, "I already use IBL a great deal" (p. 827) on a Likert scale from 1 – Strongly Disagree to 4 – Strongly Agree.

This literature review highlights that an investigation of the inquiry-oriented instructional practices of mathematics instructors has two main gaps: (1) it has not been done on widespread scale at the college level, and (2) that IOI is happening is often taken for granted without much critical thought as to what counts as IOI. That is, past studies assume that IOI exists as a recognizable, identifiable way to teach. As such, I next review the literature on IOI and organize them into a framework that can be used to build an instrument that can be used at the college level.

Organization of IOI

There have been attempts to organize the components of IOI or IBL. I mention two that have appeared recently in the literature. In one of them, Kuster, Johnson, Keene, and Andrews-Larson (2017) offered a four-principle characterization of IOI: "generating student ways of reasoning, building on student contributions, developing a shared understanding, and connecting to standard mathematical language and notation" (p. 13). In this conceptualization, the instructor acts as a broker between the field's established notion of mathematics and the students attempting to access it. In the other, Laursen and Rasmussen (2019) pose that there exist four pillars of inquiry-based mathematics instruction. These pillars include two student behaviors – that "students engage deeply with coherent and meaningful mathematical tasks and collaboratively process mathematical ideas" (p. 138), and two instructor behaviors – that "instructors inquire into student thinking and foster equity in their design and facilitation choices" (p. 138). The two student-focused goals originate in IBL work from Laursen and colleagues (2014) while the third goal of inquiring into student thinking originated in IOI work

(Rasmussen & Kwon, 2007). Laursen and Rasmussen (2019) explain that the fourth goal is somewhat implicit in the first three, as the first three pillars contribute to fostering equity in the classroom. However, the fourth pillar additionally raises awareness about attending to what students may be thinking or feeling.

I pose that conceptualizations of IBL like these are useful for unpacking the intent of instructors or making broad calls for research agendas – but they are not helpful for understanding how intent translates to action in the classroom. Such organization into broad principles is useful but still needs to be operationalized. As is, the pillars do not lend themselves well for studying how IOI or IBL is currently taught in part because the categories are not well-defined. Many actions could satisfy multiple principles, and some of the principles would not be apparent from action alone. With this framework, IOI could not be studied using methods like classroom observations without work to develop a lower-inference coding system. For example, Kuster and colleagues (2017) write that one way of generating student ways of reasoning in action is to explicitly ask students to share their approaches with the goal of finding out how they arrived at their answers. But explicitly asking students to share their approaches could easily also be part of building on student contributions or including peers to building a shared understanding of how the student did something. An instructor could have asked a student to share their approach in an effort to build that student's confidence or in order to correct their solution – it would be difficult to reliably assess intent. Kuster and colleagues' (2017) four principles could be useful for an instructor to understand the values behind his or her instructional decisions, but accurately recalling intent behind an action and assessing whether it is accurate is not easy. I use the instructional triangle to offer a framework to organize inquiry-oriented practices in a more tangible manner than the four principles or pillars. The forthcoming framework is composed of specific practices (e.g., including open problems) rather than the values motivating the actions.

Using The Instructional Triangle to Frame Inquiry-Oriented Practices

The instructional triangle (Cohen, Raudenbush, & Ball, 2003; Ball & Forzani, 2007) is a useful frame to organize inquiry-oriented instructional practices (Figure 2.1). They pose that instruction is not just what teachers "do, say, or think", but rather it is what they "do, say, or think with learners, concerning content, in particular organizations and other environments, in time" (p. 124). The instructional triangle is composed of the interactions between the teacher, students, the content, and the environment surrounding all three. In contrast with the four pillars, it provides a clearer division of possible components of IOI. Other scholars have recognized this triad (e.g., Hawkins, 1967/1974; Henderson, 1963; Schwab, 1978), but the instructional triangle proposed by Cohen, Raudenbush and Ball is unique in that it captures the dynamic between all three vertices and arrows between them over time, as opposed to using the frame to then focus primarily on only one vertex (Ball & Forzani, 2007). These relationships can be used to organize the many different conceptions of what IOI entails.

This framework builds off work done by Mesa, Shultz, and Jackson (2019), who arrived at a framework that closely resembled the instructional triangle. In that paper, they noticed in data from interviews that instructors' descriptions of IBL varied on three axes: the amount of discovery the students were tasked with, the amount of student-to-student peer engagement, and the amount of information provided by the instructor. Their study gives a specific example of the generalized framework of the instructional triangle. That is, each of their axes corresponds directly with relationships on the instructional triangle. The amount of discovery corresponds with the students' interaction with the content, the peer engagement corresponds with the student-student interactions, and the amount of information provided by the instructor corresponds with the student-instructor relationship. The framework offered by Mesa and colleagues (2019) gives support to the hypothesis that the instructional triangle is a valid, general way to capture some of the defining characteristics of different versions of IOI. The use of the instructional triangle not only gives an organization of the inquiry-oriented practices, it also motivates a closer consideration of the contexts where the practices are occurring. An added element of the triangle not featured in the framework by Mesa and colleagues (2019) is that it posits that the interactions occur within and surrounded by an environment. In the section on instructor beliefs, I show how considering the environment is a key to conceptualizing instructor decision-making.

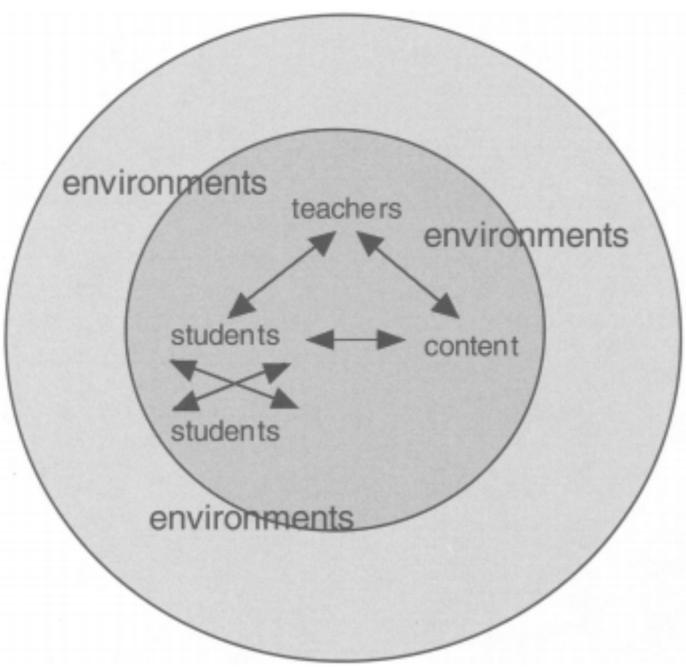


Figure 2.1: Instructional triangle from Cohen, Raudenbush, & Ball (2003), p. 124. Copyright 2003 by American Educational Research Association.

To organize the literature on IOI, I consider how the triangle organizes four possible dimensions of instructor decision-making: The instructor can choose the amount and nature of information to deliver (or withhold) from the students (teacher-student), the instructor can choose how to let the students interact and engage with each other so that they can share resources (student-student), the instructor can choose how the students are exposed to and interact directly with the content (student-content), and the instructor can choose how to interact with the content him or herself (teacher-content), for example, studying the content on their own or creating resources for students. I use the instructional triangle to review what IOI entails to mathematicians and education researchers thus far.

IOI Enacted Between the Student and Content

This section first draws from the science education inquiry literature to investigate what is meant by inquiry, in terms of students' interaction with the discipline-specific content. Then it turns to the work of mathematicians, and what the IOI literature recommends being transferred from the work of doing authentic mathematics to the classroom.

Inquiry-oriented learning has been valued among different disciplines and has a particularly strong literature base in science education. Across disciplines, a principal goal of inquiry-based instruction is to teach students to think and act like scholars (Gonzalez, 2013). The National Research Council released the *Science Education Standards* in 1996 with a central focus on the teaching and learning of inquiry that enabled scientific investigation, which prompted the release of *Inquiry and the National Science Education Standards: A Guide for Teaching and Learning* (National Research Council, 2000) and spurred much of the recent research. As mentioned in the first chapter, Pedaste and colleagues (2015) reviewed almost 60 articles on the use of IBL in science classrooms and offered the following definition: "Inquiry-based learning is an educational strategy in which students follow methods and practices similar to those of professional scientists in order to construct knowledge" (p. 48). Education theorists have also stressed the importance of students' school experiences aligning with the disciplinary practices of scholars across any subject (Bruner, 1960; Dewey, 1902; Schwab, 1978).

For mathematics, inquiry-oriented learning similarly involves following the methods and practices of mathematicians (Yoshinobu & Jones, 2012) or getting students to engage in "authentic mathematical activity" (Johnson et al., 2013). Pedaste and colleagues (2015) continue, "It can be defined as a process of discovering new causal relations, with the learner formulating hypotheses and testing them by conducting experiments and/or making observations (…) Inquiry-based learning emphasizes active participation and learner's responsibility for discovering knowledge that is new to the learner" (p. 48). Though mathematical inquiry does not involve conducting experiments in the same sense as in science, it is worth asking what might be the equivalent of experiments in inquiry-based mathematics learning.

The work of mathematicians involves contributing ideas, struggling with definitions, experimenting with examples, proposing conjectures, propositions, and theorems, and providing proofs and arguments for those claims. Lakatos (1976) wrote in *Proofs and Refutations* about an imaginary classroom dialogue surrounding the problem of finding a relation between the number of vertices, edges, and faces of polyhedra. As the class progresses, they consider examples and counterexamples, revise the definition of polyhedral, and construct various conjectures and proofs. The process is anything but linear and conjectures are continually revised as students encounter new evidence and arguments posed by each other. This exploratory discovery has been cited by mathematics educators to explain what it means to think mathematically (e.g., Schoenfeld, 1992). Among the several seminal works on mathematical problem solving Pólya wrote, a section in *Mathematical Discovery* is titled, "Observe, generalize, prove, and prove again" (Pólya, 1962, p. 76). That text shows by example that doing mathematics not only involves solving problems, but also formulating hypotheses from observations and *problem-posing* (Silver, 1994, 1997).

A common practice in IOI is presenting students with a sequence of problems that scaffold the students' discovery of some intended content (Yoshinobu & Jones, 2012). For example, Wawro, Rasmussen, Zandieh, Sweeney, and Larson (2012) designed a sequence of tasks known as The Magic Carpet Ride in a linear algebra class. The problems supported the reinvention of formal definitions for span, linear dependence, and linear independence. Students often struggle using formal definitions (Edwards & Ward, 2004; Vinner, 1991; Wawro, Sweeney, & Rabin, 2011), and The Magic Carpet Ride activity provided students with the opportunity to formulate, defend and utilize their definitions. However, there are varying degrees to which discovery or inquiry can be the centerpiece of inquiry-oriented learning. For example, Mahavier (1999), a practitioner of the Moore Method, does not consider it reasonable for a student to completely find out enough on his or her own to cover the material listed on the syllabus. He claims that *cooperative learning* might be a better term than *discovery learning* to capture the instruction taking place. Likewise, Yoshinobu (2012) describes different levels of IBL, and among the work done across levels, full discovery does not appear until the highest level. Though referring to science rather than mathematics instruction, Banchi and Bell (2008) similarly perceive the levels of inquiry as a continuum, depending on how much information and guidance is provided for the students. None of those articles assume that the highest level of inquiry with full discovery is the ideal; indeed, they suggest that there are situations where lower levels of discovery are likely more appropriate for a given environment of instruction.

Whether or not discovery is emphasized, providing students with opportunities to solve problems has been considered an aspect of IOI. Mathematicians have often recognized the centrality of problem-solving to their profession (Parker, 2005; Ulam, 1976; Wiener, 1956). For example, Halmos (1980) says that "problems are the heart of mathematics" (p. 524). Not only can students discover or rediscover new conjectures and claims, they can also discover mathematical methods for solving problems. Open problems, which are novel problems that can be approached multiple ways and may have more than one solution (Lubienski, 2000), can give students the chance to work mathematically and discover solutions that involve more than replicating a studied example. As part of the Inquiry-Oriented Differential Equations Project, Rasmussen and Kwon (2007) described how students spent extensive time developing different approaches instead of following a set of prescribed processes. Solving open problems requires conceptualizing situations mathematically, rather than merely figuring out what to "do when you are stuck" (Lesh & Harel, 2003, p. 160). As an example, Lesh and Yoon (2004) provided students with a quilting task that asked students to write instructions on how to make template pieces that are the exact size and shape given the photograph for any quilt. The open problem required generalizing observed patterns, providing a solution, and arguing why that solution fit every possible circumstance.

A common thread among the relationship between the student and content in inquiry-oriented learning is empowering the student to engage with the mathematics and view mathematics as a human activity (Rasmussen & Kwon, 2007). Though the goal is to give students as much engagement with the content as possible, the instructors' role to enable this is not trivial. I turn to the next relationship in the instructional triangle, between the students and teacher.

IOI Enacted Between Students and the Teacher

The extent to which the instructor facilitates and scaffolds varies widely in the literature. It is a common misunderstanding that in all versions of IOI, students are left to discover everything on their own; Hmelo-Silver, Duncan, and Chinn (2007) contend that the instructor should provide extensive scaffolding to help students learn how to handle the increased cognitive load. Mahavier (1999) suggested that it does not matter under the Moore method whether an instructor lectures a lot or a little, as long as there is regular interaction with the students. While strict lecture with the instructor dictating the lesson in the front of the classroom may be thought of as providing the least opportunities for inquiry, there are ways to make lecture more responsive. Burn, Mesa, and White (2015) used the term *interactive lecture* to refer to presenting material in an engaging way that included questions and answers. An instructor they interviewed explained that sometimes the term *lecture* is used to indicate that the teacher is at the front of the classroom, “But at the same time they’re using questioning techniques and getting feedback from students, …not like what you might picture like a big lecture hall” (p. 35). Many instructors in their study gave students time to work on problems in class. Yoshinobu (2012) similarly described the first level of IBL with the term *active lecture* and alluded with it to instructors using methods such as Think-Pair-Share[1], asking questions about concepts, or having students work out examples. Though these practices are not necessarily asking students to discover

---

[1] Think-Pair-Share is a teaching practice in which the instructor presents, has the students group together in pairs to discuss, and has them share out to the class.

mathematics that is new to them, they contribute to making the classroom experience more responsive to the students and less centered around the teacher. Thus, the first concrete IOI practice between students and teachers has to do with making full-class lectures more interactive.

The second IOI practice that emerged from the literature stems from how instructors and students interact in the classroom in situations where the instructor is not presenting at the board, such as when a student presents, or the students are working in groups. Gonzalez (2013), a practitioner of inquiry-based learning, described his role as becoming more of a "'guide on the side' than a 'sage on the stage'" (p. 35). Mathematics educators have described the instructor's role as facilitating and guiding student discussion to develop the mathematical sophistication of student thinking (Johnson et al., 2013; Wagner, Speer, & Rosa, 2007; Wawro et al., 2012; Yoshinobu & Jones, 2012). IOI aims to utilize student ideas and justifications to fuel the mathematical progress of the class (Wawro et al., 2012). When a student is stuck, the instructor does not give the solution away, but helps by posing a question or finding a smaller problem or special case that can help them make progress on the larger problem (Yoshinobu & Jones, 2012). An instructor can also direct students towards their peers for guidance (Yoshinobu & Jones, 2012).

IOI Enacted with Students and Their Peers

In inquiry-oriented classrooms, students are often asked to present their solutions to classmates and receive feedback on their reasoning (Gonzalez, 2013; Hayward, Kogan, & Laursen, 2016; Laursen & Hassi, 2010; Yoshinobu et al., 2011) either at the front of the classroom or in small groups (Yoshinobu & Jones, 2012). In small group discussions, students often work on problems together, while during presentations, one student leads the class in finding a proof or solution and other students can comment or ask questions. Instructors sometimes assign roles to certain students to help them work together, and keep everyone engaged (Gonzalez, 2013). The Moore Method historically prohibited collaboration or conferring between students (Parker, 2005), but many interpretations of inquiry-oriented learning actively encourage it (Gonzalez, 2013; Yoshinobu & Jones, 2012). Exchange of arguments is an important part of the work of doing mathematics, as Lakatos (1976) illustrated. Burton (2004) found in a study of mathematicians that they often work collaboratively and shape ideas through collaboration in verbal or technologically mediated exchanges. Practitioners like Renz (1999) posit that when students interact, they gain motivation from their peers to check their own work carefully and present their ideas clearly.

There are instructional decisions that relate to student-student interactions at a smaller grain size than the group or presentation format. For example, in his account of how he promotes inquiry-based learning, Mahavier (1999) wrote about deciding whether to let students help the presenter when he or she is stuck. He said he decides on an individual basis, depending on whether he thinks the student would welcome help or not. I do not review all the literature on group work because such reviews exist (e.g., Davies, 2009) and is not the focus of this study. But to name a few features relevant to inquiry-oriented learning during group work, instructors must choose when to let groups struggle on their own, when to redirect questions to group members, and when or how to validate members of groups that are on the right track. These decisions about student-student interactions are all part of this dimension of the instructional triangle relevant to IOI.

IOI Enacted Between Instructors and the Content

Instructors engage with the mathematical content to design and choose the inquiry-oriented problems or activities for their students (Gonzalez, 2013). An instructor's mathematical content knowledge as well as their *pedagogical content knowledge* (as defined by Shulman, 1986) are prerequisites to this work (Wagner et al., 2007). Mahavier (1999) wrote that Moore was a master at understanding his students' capabilities and individually assigning them problems that were challenging enough to instill perseverance and pride for a student, but not so challenging that the student would grow discouraged and give up. Chalice (1995), an instructor and advocate of the Modified Moore Method, suggests that instructors can hand out a list of problems that students could choose to work on, and mark problems with asterisks to indicate difficulty level. Doing so would not only require knowledge of how each problem was solved, but also knowledge of what would be more challenging to students. Instructors also need to know the content well enough to recognize when students understand something but are not giving a traditional proof or solution or using an alternate representation.

Though materials that support IBL are increasingly available such as through *The Journal of Inquiry Based Learning in Mathematics* and instructors often share materials with their colleagues (Mesa, Shultz, & Jackson, 2019), there are less readily available resources for IOI than for traditional lecturing. The instructor not only needs to know the content well enough to tell it to the students, but also needs to know how to guide the students to that knowledge. Instructors trying to implement inquiry for their students may choose to design their own tasks that scaffold students' discovery or understanding content of their own (e.g., Gonzalez, 2013). It is common for instructors to report spending more time preparing for a class where they use IOI than a class where they are not (Mesa, Shultz, & Jackson, 2019). To help understand whether difficulties implementing IOI stem more from issues like time constraints or more from deeper-rooted opinions, I turn to the literature on instructor beliefs.

Instructor Beliefs

There is a vast and growing body of literature on teachers' beliefs. I review the literature on connection between beliefs and practice, and contrast with studies that claim beliefs often contradict practice. The latter studies motivate the theoretical framework, because I posit that there are rational justifications why instructors do not always enact their beliefs. *Beliefs* have been defined as "psychologically held understandings, premises, or propositions about the world that are thought to be true," (Philipp, 2007, p. 258) and conceived by education researchers as "a filter through which new phenomena are interpreted" for instructors (Pajares, 1992, p. 325). Mathematics education researchers have identified beliefs as having a fundamental role in explaining mathematics teaching (Pajares, 1992; Phillip, 2007; Pintrich, 1990), and it is a construct that may help begin to explain why some instructors use inquiry-oriented methods to teach and some do not. Researchers have been interested in beliefs because of their hypothesized central role in guiding judgements and decision-making (Bandura, 1986; Dewey, 1933). Literature has suggested that teacher beliefs about mathematics have an influence on students' opportunities to learn (Gellert, 2000; Speer, 2001; Sztajn, 2003) and instructional improvement (Lewis, Fischman, & Riggs, 2015). Teacher beliefs is a broad construct that can include a wide range of categories (e.g., beliefs about the role of the teacher, about student ability, about the nature of mathematics), and for this study I focus on teacher beliefs about mathematics teaching and learning. The idea being that if the object of study is the use of inquiry-oriented practices, then a logical predictor would be beliefs about whether or not those types of teaching practices are productive.

There exist a wide range of schemes in the field to categorize mathematics teacher beliefs about teaching and learning (Speer, 2005). All schemes on teacher beliefs about teaching and learning mentioned by Philipp (2007) in his handbook chapter were theorized and studied at the K-12 level. One particularly influential scheme comes from Kuhs and Ball (1986) who identified four principal categories: (1) *learner-focused* beliefs that view mathematics learning as an active, constructive process, (2) *content-focused* beliefs with an emphasis on *understanding* that view mathematics teaching as developing the students' conceptual understanding of the intended content, (3) *content-focused* beliefs with an emphasis on *performance* that view mathematics teaching as developing students' ability to carry out mathematical procedures and (4) *classroom-focused* beliefs that emphasize replicating effective classroom strategies. I would imagine that there would be a direct relationship between the beliefs instructors hold as defined by these categories and the frequency of use of inquiry-oriented instructional practices. For example, I would hypothesize that instructors with the strongest learner-focused beliefs would use IOI most often.

There is some evidence to believe that instructor's beliefs on teaching and learning are indicators of their pedagogical practice. Various researchers have studied whether such beliefs align with the use of IOI. Wilkins (2008) found that beliefs about the effectiveness of inquiry-oriented learning mediated the relationship between content knowledge, attitudes and inquiry-oriented classroom practice in a study of 481 elementary mathematics teachers. He found that, among all variables (content knowledge, attitude, instructional beliefs, instructional practice, years of experience, mathematics courses, and highest degree) measured in a path analysis, the relationship between instructional beliefs and instructional practice was strongest. Stipek, Givvin, Salmon and MacGyvers (2001) found substantial coherence between the beliefs about the nature of mathematics and mathematics learning and the observed practices of 21 elementary teachers. Teachers that believed mathematics was more about using procedures to solve problems than a tool for thought and focused on correct solutions instead of understanding were more likely to be observed emphasizing student performance and speed, and minimizing emphasis on effort and understanding. Other studies, however, have shown inconsistencies between beliefs and practice.

Contradictions Between Beliefs and Practice

Though links have been found between teacher beliefs and teacher practice, researchers have also found inconsistencies between beliefs and practice. For example, Wijaya, van den Heuvel-Panhuizen, and Doorman (2015) found that junior high school teachers in Indonesia reported in questionnaires that they found context-based tasks beneficial for students, but observations revealed the teachers did not implement them. Raymond (1997) followed an elementary teacher and found that her practices were consistent with her beliefs about mathematics content, but not about mathematics pedagogy. DeFranco and McGivney-Burelle (2001) observed 22 graduate teaching assistants (GTAs) in a mathematics department throughout enrollment in a mathematics pedagogy course and found that, though their written journal reflections demonstrated that they wrote about adopting a new set of beliefs, their classroom practices did not change. They hypothesize that these new beliefs did not change the practice of teachers partly because their students resist practices that are different from cultural norms they have been used to with other teachers.

Ernest (1991) recognized this influence of social expectations as a constraint on enacting beliefs, including the expectations of "students, their parents, fellow teachers, parents, and superiors" (p. 290). Cooney (1985) found that a teacher's idealistic beliefs around problem solving that had been expressed during preservice training conflicted with the reality of the classroom once he began teaching. The teacher realized he did not have time to engage students in "real problems," and that students were unreceptive to his initial teaching style because it conflicted with students' expectations (p. 330).

Such inconsistencies have been explained by researchers as due to other non-mathematical beliefs or contextual factors (Philipp, 2007). Instructor decisions are shaped by the way that they are situated in environments, as depicted by the instructional triangle. Beliefs are often studied as a feature of teachers' individual cognition (Speer, 2005), but "it seems that actions and beliefs are shaped by the conditions of classrooms and teacher decision stem more from the social practices which frame teaching than the cognitive structures and beliefs of individual teachers" (Hoyles, 1992, p. 37). Sztajn (2003) illustrated this phenomenon by studying two teachers with similar beliefs about mathematics but differences in instruction. The two teachers taught in separate contexts, and Sztajn (2003) could only explain the differences in instruction once she took into account the teachers' broader beliefs about children, society, and education in the contexts in which they were teaching. Skott (2001) studied a teacher who believed that mathematics learning happened best when he could engage students in an unobtrusive manner but was prevented from always enacting this belief due to conflicting priorities such as building students' confidence and managing the classroom. While he wanted to support students to work and investigate mathematics independently, such as by not offering too much scaffolding and not evaluating student responses by letting them find out if their suggestions are correct on their own, he sometimes used direct instruction due to the classroom environment. On a much bigger scale, Engeln and colleagues (2013) found that while the 917 teachers from 12 countries all had, on average, positive attitudes towards IBL, the implementation varied more widely than attitudes, country to country. This suggests that the implementation was less dependent on individual beliefs than it was on contextual and cultural factors.

Cross Francis (2015) argued that the alleged contradictions can arise because researchers entered the field investigating a certain set of beliefs and associating those with the teacher actions, overlooking other non-mathematical beliefs or factors that might have influenced those actions. The beliefs that researchers claim to be held by the teachers are subject to the methods the researcher has used to collect and analyze that data (Speer, 2005). Hoyles (1992) argued that the decontextualized beliefs that researchers use should be replaced with situated beliefs. She wrote that, "once the embedded nature of beliefs is recognised, it is self-evident that an individual can hold multiple (even contradictory) beliefs" (p. 40). Barkatsas and Malone (2005) investigated the beliefs of 466 Greek secondary mathematics teachers and used a principal components analysis to find two major orientations towards mathematics teaching and learning: a contemporary-constructivist orientation (similar to inquiry-oriented) and a traditional-transmission-information processing orientation. In the same paper, a case study of a teacher with a contemporary-constructivist orientation revealed that the teacher's practice was influenced by cultural factors and social norms (Barkatsas & Malone, 2005). For example, though teachers may believe group work provides the best environment to learn mathematics, they may feel pressure to prepare students for university entrance examinations that motivates them to prioritize direct instruction. This study frames beliefs as a factor to consider, but also draws from these past results as motivation to turn to broader theories for understanding why instructors choose to teach in certain ways.

This chapter has outlined some practices in the inquiry-oriented instruction literature, and suggested that previous results give reason to believe that beliefs might not be able to entirely explain why instructors use those practices or not. I turn to a theory of decision-making that takes contexts and environments into account when providing explanations for why instructors do what they do.

Chapter 3 **Theoretical Framework**

In this section, I begin by explicitly stating my hypothesis on the relationship between beliefs and practice, and use it to introduce the theory of practical rationality. I follow that with an outline of my conceptualization of IOI.

Beliefs as Motivation for Another Construct

Beliefs are an important lens researchers in the field have used to examine how instructors view their practice and make instructional decisions. Their beliefs act as individual resources, built on a lifetime of experiences. Philipp (2007) noted that mathematics education researchers have tried two approaches to explain inconsistencies between beliefs and practice: the first is to investigate whether multiple beliefs are at stake and illustrate how a teacher is prioritizing some beliefs over others, while the second is to examine if some aspect of the teacher's perspective on his or her practice can explain the differences. I prefer the second approach because the construct of beliefs alone, even many different types of beliefs, ignores the environmental and social contexts that teachers work in.

The literature reviewed illustrated that beliefs sometimes are not realized in practice due to environmental factors. For instance, Sztajn (2003) could only explain the differences in instruction once she considered the teachers' broader beliefs about children, society, and education in the contexts in which they were teaching. I take the view that teachers are rational beings and agree with Leatham's (2006) notion that "teachers are inherently sensible rather than inconsistent beings" (p. 92). Instead of concluding that teachers are acting irrationally or out of laziness, I argue that beliefs alone cannot account for the complexity of teaching and the environment that they teach in. Philipp (2007) urged mathematics education researchers to see inconsistences between beliefs and practices as problems for researchers, rather than the teachers, to solve. That is, researchers need to alter and improve their theories about what is occurring to make the inconsistency actually consistent with predictions. In order for the field of mathematics education to have theories with predictive validity for anticipating instructors' decisions, other theories of teaching that incorporate more than beliefs ought to be used that can account for the practical conditions that instructors have to consider. I draw from a theory of teacher decision-making to provide a more holistic and socially-situated lens for understanding pedagogical practices.

Practical Rationality

Herbst and Chazan (2011, 2012) proposed the framework of *practical rationality* to describe the sources of justifications that practicing teachers make decisions based on both social and individual resources. This framework is useful to account for the environment in which instruction (the relationships among teachers, students, and content) exists. This theory of decision-making provides a lens for how teachers justify and rationalize their decisions, rather than for explaining the cognitive processes involved in making the decision. Various researchers have agreed that studying beliefs alone is inadequate, and that they need to be contextualized in situations (Barkatsas & Malone, 2005; Cross Francis, 2015; Ernest, 1991; Hoyles, 1992). The theory complements studies of only individual resources such as beliefs or knowledge by also considering characteristics of the environment, thus aligning with Cohen, Raudenbush, and Ball's (2003) conception of the instructional triangle: "interactions among teachers and students around content, in environments" (p. 122).

Herbst and Chazan (2011) conceptualized the environment as a key element of instructional systems. Instruction draws from environments such as policies or pedagogical practices required by the institution, expectations of the students or parents, and priorities of local districts or state agencies (Cohen et al., 2003). Within these environments, instructors must negotiate tasks or series of tasks to engage students with the mathematical content being studied. Managing that engagement is complex because the instructor must both deploy mathematical content in the form of work that students can do while interpreting that work on behalf of the content being studied (Herbst & Chazan, 2011). The interaction around mathematics between instructor and students is thus mediated by more than the knowledge and beliefs of the teacher; it is also mediated by the work contexts that a teacher can create for students in institutional environments. Once students and instructors have interacted in a given environment, they develop expectations for the interactions with each other. That is, once students and teachers have interacted in the setting of schooling, they carry expectations for what those interactions will continue to look like.

The framework of practical rationality is not only valuable for its capacity to account for the environment, it also considers the social interactions that shape behavior around the work of mathematics instruction. Practical rationality is rooted in the notion of the *didactical contract* (Brousseau, 1997; Herbst, 2003), which can be described as a set of norms that regulate the work between instructors and students around the transaction of mathematical content (Brousseau, 1997). These norms around transacting the knowledge at stake are negotiated between the students and teacher and can be tacitly or implicitly imposed (Brousseau & Warfield, 2003). The negotiations result in a set of usually tacit norms that guide the work of teaching mathematical content. For example, students may expect the problems assigned by the instructor to be solvable in a short amount of time (Schoenfeld, 1988). And if problems assigned by the instructor take much longer, students may sanction the instructor's behavior by complaining. The instructor can give simpler problems, offer sanctions in the way of low grades for incomplete work, or provide some other compromise that establishes new expectations for teachers and students to follow.

The didactical contract is relevant to conceptualizing instructors' decisions to use IOI because enacting inquiry-oriented practices often requires departing from the norms usually associated with the college mathematics environment. Herbst and Miyakawa (2008) named three types of norms that describe the default way knowledge is transacted in the classroom: *division of labor norms*, for who needs to do what, *temporal norms*, how long and in what order things are done, and *exchange norms*, what needs to be done and how it should be done. I use these categories to hypothesize about the work instructors must do to renegotiate these norms for the work of teaching with inquiry. In the category of division of labor norms, consider the task of introducing a new mathematical definition to the class. Under the didactical contract that exists in most undergraduate mathematics classes, the instructor would be considered responsible for introducing this new knowledge and clarifying how to use it. In an inquiry-oriented classroom, the students may be devolved responsibility for some or all of this work. In terms of temporal norms, the introduction of a new concept might take longer if an instructor is aiming for students' deeper understanding by having students discover it by themselves. Topics might shift from being all "covered" during class to also expecting students to engage with new material on their own outside of class. As an example of an exchange norm, what sorts of claims require proof and what proof consists of might differ in a class where an instructor is trying to emphasize inquiry-oriented practices. In a calculus classroom following the usual didactical contract, instructors might hand-wave some proofs of theorems and see the ultimate goal being that students can complete the procedural homework problems. In a classroom in which the instructor is attempting to incorporate inquiry, they might delve deeper into some proofs in hopes that students gain some conceptual understanding of the process. Or conversely, if an instructor is having students come up with all the proofs from scratch, there might be different criteria for what counts as proof (e.g., not as formal or polished as the proofs in a published text).

For any form of IOI, the didactical contract needs to be renegotiated between the instructor and students, and the instructor is arguably the person with the power and motivation to catalyze those negotiations. Given that in the didactical contract the instructor is the party responsible for transacting the new knowledge, and that some instructors believe that the best way to do so is by having students engage more actively with the content using inquiry-oriented practices, the instructor then has to do the initial work to breach existing norms and redefine the contract. This work is not simple and helps to explain the larger story of why sometimes instructor's beliefs do not align with their practice. In addition to the negotiation between instructors and students, the didactical contract involves the environment and an educational system in establishing the transaction of mathematical content (Brousseau & Warfield, 2003). The teachers' obligations to the rest of society and the school system are aspects of this transaction (Brousseau & Warfield, 2003). The framework of practical rationality calls these professional obligations.

Professional obligations

Herbst and Chazan (2011) hypothesize that a source of justifications for decisions in institutional environments where mathematics instructors work come from four *professional obligations* that mathematics teachers must respond to as professionals: towards creating a socially and culturally appropriate environment for students to share space and resources (interpersonal), representing the discipline of mathematics appropriately (disciplinary), respecting institutions such as the school, department, or state in matters including curriculum, assessment, and policy (institutional), and to treating individual students as persons with unique assets and needs (individual). The obligations correspond to various stakeholders in the educational system, namely society, knowledge, organization, and the client, respectively (Chazan, Herbst, & Clark, 2016; Herbst & Chazan, 2012). These obligations can also be thought of as shared beliefs sanctioned by the profession of teaching mathematics about the responsibilities which professionals may or may not choose to prioritize (Chazan, Herbst, & Clark, 2016). They are what teachers would typically use implicitly to justify their practices to stakeholders. The theory provides a lens to address the "multiple and sometimes conflicting educational priorities" (Skott, 2001, p. 18) that arise in the practice of teaching.

The mathematics education literature provides a basis for seeing these obligations as a major aspect of mathematics instruction. Webel and Platt (2015) found that the way two high school mathematics teachers recognized professional obligations impeded their ability to actualize their goals in their instructional practice. In their study, the teachers' decisions were guided by disciplinary and individual obligations that inhibited them to make the desired changes. Their finding supports my hypothesis that professional obligations can sometimes impede the realization of what instructors believe would be optimal for student learning. Lande and Mesa (2016) found that professional obligations played an important role in community college instructor's decision-making and used the obligations to frame comparisons between part-time and full-time faculty regarding their sense of agency.

There are studies that allude to professional obligations without naming them directly. For instance, Lampert (1985) described her own struggles making decisions as a teacher, saying that, “The contradictions between the goals I am expected to accomplish thus become continuing inner struggles about how to do my job” (p. 182). Her choice of the word “expected” indicates that she is implicitly attending to professional obligations, without explicitly saying what stakeholders are doing the expecting. The next section reviews literature relevant to each professional obligation and connects the obligation to the decision to use inquiry-oriented practices.

Relationships Between Professional Obligations and IOI

The Disciplinary Obligation and IOI

There is evidence to support that there might be a relationship between professional obligations and IOI. I highlight potential links in the disciplinary, individual, interpersonal, and institutional obligations to IOI, in that order. Some researchers (e.g., Ball, 1993; Lampert, 1990) have found that what teachers deem to be appropriate actions are based on considerations concerning the discipline of mathematics (the disciplinary obligation). Education theorists like Buner (1960), Schwab (1961/1974), and Wineburg (1989) advocated that the schooling system more closely resemble the activity in the disciplines being taught. Even if an instructor holds beliefs not conducive to including inquiry, they might include some inquiry-oriented practices regardless of making the classroom more akin to the work of mathematicians. For example, an instructor might believe that students do not necessarily have to experience struggle in order to learn the mathematical content, but still include opportunities to solve open problems and make novel conjectures (activities that often can be frustrating for students) because they want the course content to more authentically reflect the practice of mathematicians. Conversely, there may be instructors who tend to use practices that are not inquiry-oriented to uphold other aspects of the disciplinary obligation. An instructor could desire to present a theorem or proof in the most elegant way with canonical notation, and thereby not wish to let the students discover it themselves. As an example, Lew and colleagues (2016) documented a professor respected for his teaching, who presented proofs in lecture format with the intent of allowing students to see the formal product written on the board while using his speech to explain the process of constructing it. He chose a style of instruction without inquiry to portray two aspects of a fundamental skill in the discipline.

Inquiry-oriented instruction has been defined as giving the students the opportunities to engage in practices similar to professionals in their field (Pedaste et al., 2015). Instructors may have reasons tied to the professional obligations that they do not want their students to replicate how mathematicians do their work with complete inquiry. Ball (1993) explained how in her own practice she understands the value of creating an authentic disciplinary experience for her students, but also faces some dilemmas inherent from the nature of mathematics as a field. First, mathematicians tend to work by themselves on one obscure problem for long periods of time. In a classroom, teachers must focus on helping all students learn the content in the curriculum together, simultaneously. Second, the field of mathematics is competitive and mathematicians strive to improve their prestige. Ball (1993) did not want these elements of authenticity to exist in her elementary classroom. On the opposite end of the spectrum, the Moore Method had a way of fostering competition and isolation (Cohen, 1982). Views of how close the classroom should be to the discipline certainly shaped the practices of Moore and Ball. For example, competition and public recognition was a driving force for the students in Moore's courses (Jones, 1977), while Ball (1993) hoped to foster social norms that inclusively respected the reasoning of all students. This literature contributes to my hypothesis that the disciplinary obligation plays a significant role in both if and how various inquiry-oriented practices are implemented.

The Individual Obligation and IOI

There are potential links between the individual obligation and inquiry-oriented practice. Studies on beliefs found that attending to students' individual needs (the individual obligation) mediated beliefs about mathematics (Sztajn, 2003; Skott, 2001). Skott (2001), for example, related the case of a teacher who

> consistently pointed to the need to think of the children in broader terms than those related to mathematics. In general, and in relation to individual children, he showed a lot of sensitivity to many other aspects of his students' lives, in particular to their development of self-confidence, both mathematically and otherwise. (p. 16)

There is evidence that the individual obligation varies according to culture. For example, Cai and Wang (2016) compared "distinguished" mathematics teachers from China and the United States, finding that teachers from the United States placed more emphasis on being sensitive to students' needs, having a good sense of humor, and facilitating individual student participation, while teachers from China placed less emphasis on individuals in favor of more emphasis on broad and deep content coverage and coherent lesson plans.

There are many ways the individual obligation could influence the use of some inquiry-oriented practices. For example, if an instructor wanted to foster the confidence in students to be able to solve a type of mathematical problem on their own, this might make the teacher more inclined to use some inquiry-oriented practices, such as by giving students time during class to struggle with and complete the first couple practice problems individually or in groups. Having this activity occur in the classroom allows the instructor to scaffold students' activity rather than risk the students struggling alone without the ability to get immediate guidance from the instructor or peers in the classroom. Then students could experience some amount of success before embarking on the struggle without help as readily available. Conversely, if a teacher wanted to help students feel comfortable with a certain concept on an upcoming exam, the teacher might present students with a solution and with many exercises that could be solved in a similar way. For example, community college instructors are known to prioritize giving students a sense of achievement and self-confidence at the cost of not giving novel, challenging mathematical tasks (Mesa, Celis, & Lande, 2014). Thus, those instructors might be less inclined to use inquiry-oriented practices on account of fostering the individual students' feelings of competency with the mathematics.

An instructor with a strong sense of the individual obligation might be more inclined to use inquiry-oriented practices such as waiting before telling how to do something. Mahavier (1999) contended that a main reason people who try to teach using inquiry-oriented practices find themselves without success is lack of patience for individual students. He continued that not only do instructors not have patience for the slower coverage of content, but they also struggle to accept inelegant student work (e.g., proofs that are longer than necessary) and to let students figure things out for themselves. More recognition of the individual students' experiences might overcome whatever discomfort is felt from waiting for students to respond. It is clear based on the literature that there is likely a relationship between instructors' recognition of the individual obligation and their use of IOI, though which direction this relationship tends to occur is less certain. A similar relationship – existent but not clear in which direction – can be found between the interpersonal obligation and IOI.

The Interpersonal Obligation and IOI

The existing literature related to the interpersonal obligation implicates the instructor using some components of inquiry-oriented practice. Some of the ways the interpersonal obligation has been recognized by education researchers are in terms of the importance of social networks for learners (e.g., Gholson & Martin, 2014), facilitating an environment predisposed to quality discourse (Walshaw & Anthony, 2008) and group work (e.g., Wawro et al., 2012), and attending to issues of equity realized in the classroom (e.g. Gholson & Martin, 2014; Sztajn, 2003). Gutiérrez (2013) called for increased focus on the sociopolitical issues present in the mathematics such as those raised by critical race theory and post-structuralism, with attention to the positioning of students in terms of identity and power. When instructors navigate such issues in an attempt to provide a more equitable environment, they are attending to the interpersonal environment in the classroom. A teacher attending to this obligation might be more inclined to use some inquiry-oriented practices such as using group work and reflecting authority back to the mathematical content and to the voices of other students, instead of telling students the correct way of doing things. Conversely, an instructor who is aware but unconfident about addressing issues in their classroom related to fostering an equitable, productive social environment might refrain from using IOI for fear of doing so in an unequitable way. There are other ways recognition of the interpersonal obligation can be interpreted, such as in matters of fairness or distribution of resources. An instructor hoping to give students fair preparation for college entrance exams might give students less open problems and other novel tasks in order to give them more time to practice the sorts of multiple choice problems that would appear on the exam. Recognition of the institutional obligation might similarly be inversely related to the amount of inquiry-oriented practices an instructor uses.

The Institutional Obligation and IOI

Finally, recognition of the institutional obligation can constrain or enable an instructor to use inquiry-oriented practices. The institutional obligation has been viewed as a constraint for the realization of teacher beliefs (Cross Francis, 2015; McGivney-Burelle, DeFranco, Visonhaler, & Santucci, 2001; Zeichner & Tabachnick, 1985). Instructors may be limited by institutional or departmental policies. Courses with more than one section that are taught by many instructors are often taught as *coordinated courses,* where instructors work together to ensure all students are taught "consistent core material" (Rasmussen & Ellis, 2015, p. 107). Those authors say that coordinated courses tend to be lower division courses with many sections (e.g., calculus) and typically involve features like common syllabi, exams, homework, and textbook. These features have the potential to limit what the instructors feel capable of implementing. Depending on what the common features are, they could constrain or enable the instructors to implement inquiry-oriented practices. Even if the course is not coordinated, there are basic factors such as class size, exams, content, and limited class time which can constrain or enable the extent to which instructors feel they can alter how they instruct (McDuffie & Graeber, 2003).

One common complaint with implementing IOI in courses with some level of coordination is what Yoshinobu and Jones (2012) refer to as "the coverage issue" (Johnson et al., 2013; McDuffie & Graeber, 2003; Wu, 1999). Instructors are often concerned that IOI takes too much time because topics are taught in more depth and students are given more time to engage with the content and each other (Yoshinobu & Jones, 2012). One mathematician who was implementing inquiry-oriented learning in an abstract algebra course stated, "I would say the coverage issue, in my mind, is the main concern I would have with group work. …When you have the material reinvented, in my mind it is going to slow down things" (quoted in Johnson et al., 2013, p. 746). There is a finite amount of time, and many courses have a set curriculum. Instructors feel the need to teach the topics that will be on the exam or necessary for the next course in the sequence (Johnson et al., 2013). However, there are mathematicians that use inquiry-oriented practices and claim that it is possible to cover as much, if not more, content than in a more traditional course (e.g., Chalice, 1995; Gonzalez, 2013; Yoshinobu & Jones, 2012). Or as another mathematician interviewed by Johnson et al. (2012) put it, "There are a lot more things you could say out loud in a semester, if you don't let students work through the ideas and they don't learn any of them," (Johnson et al., 2013, p. 751) concluding that students cover more in an inquiry-oriented course, if "coverage" means understanding the content. Thus, a way that the didactical contract might need to be altered in IOI courses is that instructors negotiate with students to expect to be taught portions of the content in a deeper, more meaningful way. In turn, students may need to expect to cover more of the content on their own.

Even if an instructor is less worried about covering the prescribed topics, instructors often lack the time to prepare and learn how to use IOI. For example, an instructor attempting to implement IOI in a differential equations class said that during class discussions he was not sure how to guide student ideas and move them forward (Johnson et al., 2013). The time it takes to learn how to guide student ideas and to design activities that will foster productive struggle can be a deterrent from altering instructional practice; these instructors have many competing priorities such as teaching multiple classes, holding office hours, doing their own research, and participating in other departmental or institutional activities (McDuffie & Graeber, 2003). Faculty may not be motivated to invest the time because institutions incentivize accomplishments or performances other than teaching effectiveness in decisions of tenure or promotion. Walczyk, Ramsey and Zha (2007) found that among science and mathematics faculty in Louisiana, most weighted teaching effectiveness as less than 50% of considerations contributing to institutional personnel decisions. These are all institutional issues that instructors can feel obligated to attend to.

However, Herbst and Chazan (2011) caution that the institutional obligation need not always be viewed as a constraint, but rather can also facilitate certain things. For example, professional development required by the institution could enable instructors to learn new pedagogies to address instructor's conflicting beliefs about teaching and learning mathematics (McGivney-Burelle et al., 2001). The institutional obligation to have office hours could have beneficial outcomes for students. Some mathematics departments actively promote and support the use of innovative instructional practices, so for instructors at those institutions, incorporating innovative practices could be partly due to the institutional obligation. The Harvard calculus reform is an example of such a mandate, where various departments chose to require instructors to use the materials produced, which has shown benefits for students including practitioner reports of increased engagement (Lock, 1994) and performance (Kerry, 1995). In particular, the institutional obligation could directly involve implementing more inquiry-oriented practices if the university or mathematics department stresses it.

There is evidence that there is increased institutional pressure to implement more inquiry-oriented practices. The pressure on institutions has come from a variety of sources. Advisory groups like the President's Council of Advisors on Science and Technology (2012) have called for increases in STEM majors, and recommended that universities achieve this by having their instructors follow evidence-based research. Evidence-based research trends towards inquiry-oriented practices. In a metanalysis of 225 studies, Freeman and colleagues (2014) found overwhelming support for active learning over lecture to maximize student learning and course performance (based on exam performance and course grades). The consumers of higher education, for example, parents and college students, value and expect things like smaller classes. Class size as an indicator of academic quality was added as a factor in the U.S. News and World Report college rankings in the 1990s due to research connecting smaller class size to higher student engagement (Supiano, 2018). Given these institutional pressures, I could expect to see that instructors with high recognition of the institutional obligation would use inquiry-oriented practices more frequently because there is institutional encouragement within mathematics departments to make learning more active. In order to study the relationships between the professional obligations and inquiry-oriented practices, I offer a concrete conceptualization of inquiry-oriented instructional practices.

Conceptualization of inquiry-oriented instructional practices

The inquiry-oriented instructional practices that I operationalize for this study are shown in Table 3.1. They are all directly synthesized from the researcher and practitioner accounts in the literature review, organized by the components of the instructional triangle. The student-content relationship for IOI involves engaging students with the practices of mathematicians and engaging deeply with mathematical ideas. This involves presenting students with tasks where there is more work to be done than following a procedure. I call these *open problems,* or problems that either have multiple solutions or multiple nontrivial ways of arriving at a solution. Lakatos (1976) captured the practice of construing claims and arguments and refuting them to refine, improve and create new claims, which I have split into the categories *constructing* and *critiquing*. Finally, creating definitions is an important part of inquiry in mathematics.

Definition-formulating involves inventing or reinventing definitions with the benefits of both helping students learn the practice of creating definitions and better understand how to utilize definitions, such as in the Magic Carpet Ride activity for linear algebra students described in the literature review as a type of inquiry activity that engages students directly with the content.

Table 3.1: Conceptualization of inquiry-oriented instructional practices

Triangle Relationship	Constructs	Description
Student-Content	Open problems	Posing problems that either have multiple solutions or multiple nontrivial ways of arriving at a solution
	Constructing	Posing tasks that ask students to make conjectures and construct arguments
	Critiquing	Asking students to critique the reasoning of themselves and others
	Definition-formulating	Inventing or reinventing mathematical definitions
Teacher-Student	Interactive lecture	Instructing to the full class while asking for feedback from students, asking questions of students, and having students engage with the mathematics
	Hinting without telling	Guiding a student to work productively without directly telling the student a correct way to proceed
Student-Student	Group work	Creating an environment where students work together on mathematical tasks or problems
	Student Presentations	Having a student or students present completed or in-progress work to the class
Teacher-Content	Class preparation	Planning a lesson to intentionally contain opportunities to engage in inquiry-oriented learning around the content being taught

For the teacher-student relationship, two primary ways an instructor can interact with students are the activity structure of classroom interaction and how instructors respond to individual requests for help. These interactions are captured by the constructs *interactive lecture* and *hinting without telling*, respectively. Interactive lecture can range from traditional lecture in which the instructor dictates everything to the students, to more interactive lecture in which instructors are actively seeking feedback from students and asking them questions. Hinting without telling occurs when a student asks a question or claims to be stuck, usually not during full-class instruction, and the instructor decides how to guide the student. With no inquiry-orientation, an instructor could directly tell the student what to do. Other efforts to guide the student without telling direct instructions would be considered inquiry-oriented.

Student-student interactions can be fostered either through *group work* or *student presentations*. In group work, students work together, assisting each other to solve problems, prove theorems, or do whatever mathematical activity the instructor assigns. Sometimes group work can entail assigning different roles to students. During student presentations, a student (or students) present(s) work that they are currently doing or have completed at the board in front of the class. In both of these formats, students can critique, question, comment on, or further develop each other's ideas. Finally, a primary way that the teacher-content relationship is realized is by the instructor's *class preparation*. This involves how the instructor plans the lesson, such as by creating worksheets or lecture notes with the intent of fostering discovery or any of the inquiry-oriented practices identified thus far for students. The other time instructors interact with the content is during class while teaching or looking at student work. What makes those interactions with the content part of inquiry-oriented practices depends on how the instructor delivers and responds to that content, which is absorbed in the other relationships of the triangle. I summarize what I investigated given this conceptualization of IOI and framework of practical rationality with the research questions guiding the study.

Research Questions and Hypotheses

Given the gaps identified in the literature review (i.e., a lack of quantitative large-scale, systematic study of inquiry-oriented practice in college undergraduate mathematics and of beliefs and the framing and instruments to do so) and the potential relationship between professional obligations and IOI, I pose the following research questions:

1. What inquiry-oriented instructional practices are used by college mathematics

instructors?

2. What factors might predict the use of inquiry-oriented instructional practices?
   a. Is there a relationship between the beliefs of instructors and inquiry-oriented practices or patterns of inquiry-oriented practices in college mathematics instruction?
   b. If so, do professional obligations help explain the relationship between beliefs and inquiry-oriented practices and patterns of inquiry-oriented practices in college mathematics instruction? If so, how?

The first question pertains to the ultimate goal of this study: finding out what inquiry-oriented instruction is in the context of college mathematics. Given those findings, Question 2 is aimed towards understanding what might play a role in the instructional decisions to use those practices. Question 2a asks more specifically if beliefs are a factor that can predict practices, based on literature linking the two constructs. Question 2b directly addresses the beliefs literature that has noticed inconsistencies between instructor beliefs and practice by exploring how the social resource of professional obligations might help explain why this misalignment exists.

Question 2b also asks how the four professional obligations help explain the beliefs-practice relationship. Based on the literature and theory reviewed, I hypothesize that professional obligations help explain variance in the relationship between beliefs and some of inquiry-oriented practices. The way they do so likely depends on the inquiry-oriented practice at stake. For instance, the time and curriculum constraints associated with the institutional obligation might not limit the use of interactive lecture in the same way that it does the use of student presentations or opportunities to craft definitions. The next section explains how I collected data and measured these constructs.

Chapter 4 **Methods**

I used four self-report instruments to collect data for this study: INQUiry-oriented Instruction REview (INQUIRE), PRofessional Obligations Scenario Evaluation (PROSE), beliefs, and a background survey. I first explain limitations of self-reports so I can explain how each instrument is designed to address those limitations. Then I explain sample design and the strategy for data analysis.

Self-report Survey Design

Self-report instruments are one of various ways to collect data on instruction. Kennedy (1999) called different methods of data collection *approximations*, with classroom observations and achievement tests constituting first-level approximations, responses to vignettes or logs as second-level approximations, and espoused principles and practices (e.g., survey data) as third-level approximations. She found evidence of relationships between these self-reported second- and third-level approximations and student outcomes in various studies, indicating the value of self-reports despite potential validity issues. She noted methodological advantages such as enabling aggregation of data, groups comparisons, and ease of data collection.

Given the usefulness of self-report surveys for data collection, they are frequently used in large-scale studies to study teacher behavior and have yielded valuable data for researchers (e.g., in the National Assessment of Educational Progress, the Schools and Staffing Survey, the Third International Mathematics and Science study, and in Smith, Desimone, & Ueno, 2005), but have been questioned for their accuracy (Rowan, Harrison, & Hayes, 2004) and ability to reliably elicit data for the study of implicit constructs such as beliefs or values (Finch, 1987). Self-reports can be driven by social desirability (D'Onofrio, 1989) and respondents may interpret survey items differently than intended (Ross, McDougall, Hogaboam-Gray, & Sage, 2003).

Despite criticisms, Hayward, Weston, and Laursen (2018) found that self-report surveys were trustworthy when used to describe what is happening in instruction, in comparison to observation protocols designed to measure the same happenings. Similarly, Desimone, Smith, and Frisvold (2010) found that there were only small differences between student reports of teacher behavior and teacher's self-reports, that disappeared when controlling for background variables.

Researchers have documented ways to increase the validity and reliability of self-reports. Rowan and colleagues (2004) tried to improve the content validity of the responses by giving respondents definitions of the terms in the surveys, collecting responses more frequently to reduce inaccuracy due to lapses in memory, and being available to participants for frequent communication. Content validity can also be improved by consulting with experts and by generating items from credible sources (Ross et al., 2003). Before surveys are administered, researchers can use cognitive interviews to check that respondents interpret the survey items in the manner intended, which have become one of the most prominent methods for detecting issues with survey questions (Beatty & Wills, 2007; Karabenick et al., 2007). I incorporated these elements in the self-report instruments designed for this study to maintain content validity while enabling me to collect a large enough set of data for the desired analyses.

Design of the INQUIRE Instrument

All items for INQUIRE were based on the conceptualization of IOI outlined in the theoretical framework. The instrument was designed to ask questions about inquiry-oriented practice in relation to each of the relationships of the instructional triangle. Some example items can be seen in Table 4.1 (see Table 3.1 for descriptions of the constructs).

Table 4.1: Structure of the latent constructs and corresponding examples in INQUIRE

Triangle Relationship	**Constructs**	**Example Items**
Student-Content (24 items)	Open problems	How often do you task students with problems where there are multiple solutions?
	Constructing	How often do you ask students to generalize a claim?
	Critiquing	How often do you provide students with arguments for them to critique?
	Definition-formulating	How often do you ask students to revise a definition?
Teacher-Student (13 items)	Interactive lecture	During instruction, how often do you check to see if students are following your lesson?
	Hinting without telling	If a student asks you to look at his or her work, how often do you respond without evaluating whether or not it was correct?
Student-Student (16 items)	Group work	How often do you have students work together in groups?
	Student Presentations	How often do you have students present work to the class?
Teacher-Content (9 items)	Class preparation	How often do you design a sequence of problems so that students will discover something?

I tested the validity of the INQUIRE items by conducting three cognitive interviews and collecting data from seven extended questionnaires to check of questions were interpreted as intended. They were done with 10 instructors: five mathematics doctoral students, four mathematics education doctoral students, and one faculty member of a mathematics department, all from two midwestern Research I universities. All had at least three years of experience teaching at the college level. The extended survey was taken online in Qualtrics and after each closed survey item they also responded to a question that asked how they interpreted the previous question and whether anything was unclear. Other questions were interspersed throughout the survey asking how participants interpreted the response options, and if the options captured what they would like to say. Three were in-person, to pick up on nonverbal cues such as confusion or reluctance (Holstein & Gubrium, 1995) and to be able to ask reactive probe questions.

The INQUIRE items were created to maximize the accuracy of the self-reported data. Self-report instruments that are focused on one specific area of content (Stephen & Burns, 1986) and that are situated in a specific classroom context (Kennedy, 1999) yield more accurate results than those without such focus. The participants were asked to complete all questions twice – once while thinking of a lower-division course they have recently taught, and a second time while thinking of an upper-division course they have recently taught. Yoshinobu & Jones (2012) explained that while inquiry-oriented methods can be used in any mathematics course, it might look different in upper-division courses (e.g., having more opportunities to do proofs) than in lower-division courses (e.g., having more opportunities to solve open problems). I defined lower-division courses for participants as including courses such as calculus courses, introductory linear algebra, differential equations, and introductory courses like college algebra, pre-calculus or trigonometry. Examples of upper-division courses given to participants included modern algebra, analysis, topology, advanced linear algebra, combinatorics, (topics in mathematics) for future teachers, and introduction to proof. Completing the survey with a specific course in mind may have increased the precision of responses because self-reports were found to be more accurate when they focused on a specific time frame (Stephen & Burns, 1986) and when they reported specific practices over a brief instructional period (Newfield, 1980). Accordingly, each question was designed to ask about a practice in the context of a specified time frame for a particular course.

Participants responded on the following six-point scale: 1 – Never, 2 – For a few classes, 3 – Less than half the classes, 4 – More than half the classes, 5 – Every or almost every class, 6 – Multiple times per class.[2] Likert-type scales usually offer five or six response options depending on the designer's intent (DeVellis, 2016). An even number of options was chosen so that responses could be dichotomized for analysis later. Education researchers have used other response options to test frequency of behavior such a 5-point scale from 1 - Never to 5 - Once per day (Jacobson & Izsak, 2015) or dichotomous options for instructional topics covered like "a focus of instruction" or "touched on briefly" (Rowan et al., 2004). The INQUIRE instrument contains response options as specific as possible, while allowing for the variation among mathematics classes that the instructors will have taught. Other mathematics education surveys such as the TIMSS teacher questionnaire (IEA, 2014) have used similar response options for collecting data on teaching practices. I use a different style of self-report instrument to collect data on professional obligations, which are scenario-based assessments.

Design of the PROSE Instruments

The PRofessional Obligations Scenario Evaluation (PROSE) instrument (Herbst & Ko, 2018) was a four-part multimedia questionnaire that used storyboarded vignettes. Vignettes offer a method to survey less explicit constructs like attitudes (Finch, 1987; Neff, 1979). Participants may feel less threatened when responding to the hypothetical stories of third parties than when reporting on their own actions (Finch, 1987). Finch (1987) used vignettes to study beliefs and

---

[2] Some of the items were operationalized on a 5-point scale due to an editing error from the previous version of the survey. Those items included the lower-division lecture and hint but not tell items, and all the upper-division items. The 5-point scale was as follows: 1 – Never, 2 – A few classes, 3 – About half of the classes, 4 – Every or almost every class, 5 – Multiple times per class.

found it a more convincing method to elicit responses than other methods like attitude surveys that were too direct and did not offer enough contextualization. One type of survey involving vignettes that are widely used and deemed effective as self-report instruments are situational judgement tests. These tests are used to assess the decision-making of employees in a given workplace (Whetzel & McDaniel, 2009). They typically provide a scenario in the form of a vignette, and ask the participant to respond by choosing an appropriate response from a closed set of options. Through specific scenarios, they can test qualities of the participants such as their communication skills or tendency towards teamwork in the same way that survey items test constructs.

From a researcher's perspective, third-level approximations like responses to fixed scenarios, as opposed to first-level approximations like classroom observations, have the advantage of allowing the designer to control the situation and what information is provided to the respondent (Herbst, Chazan, Chen, Chieu, & Weiss, 2011). While the scenarios in vignettes are often presented in textual descriptions, the text can denaturalize and oversimplify the situation. Information that would not necessarily be available has to be made explicit in texts. Vignettes could also be presented in the form of videos as an alternative to text (e.g., Kersting, Givvin, Sotelo, & Stigler, 2010). Videos more closely resemble familiar classrooms and the actual situations that instructors would face. Unlike text, videos enable the viewer to experience the pace and rhythm of events, and how instruction occurs in conjunction with the activity and environment of the classroom (Herbst et al., 2011). Yet, using videos to study general practices can be problematic because videos of actual classrooms contain particular aspects of its setting (what Herbst et al., 2011, called *individuality*) that might distract participants from responding to the practices shown (Chazan & Herbst, 2011).

Comic strips with speech or thought bubbles or animations offer the affordances of both text and video: they allow the designer to control the content of the vignette while providing a contextualized, temporal representation of the instruction. Researchers can depict specific events even if video records of such an occurrence are not easily accessible. They have been used by mathematics education researchers to study professional knowledge (Herbst & Chazan, 2015; Moore-Russo & Viglietti, 2010), decision-making (Erickson & Herbst, 2016; Herbst, Chazan, Kosko, & Dimmel, & Erickson, 2016), and instructor professional development (Hayden, Moore-Russo, & Marino, 2012; Herbst et al. 2011; Moore-Russo & Wilsey, 2014). The storyboards and animations in the PROSE instrument were created to meet design principles: they provide enough context that participant can understand the situation but leave enough unknown so that the participant is not distracted (Finch, 1987); they should be plausible and realistic from the participant's perspective (Neff, 1979); and groups of vignettes can be built so that constructs being tested can be generalized beyond the specific circumstances in a given story (Finch, 1987).

The scenarios in the PROSE instrument were created using Depict in the Lesson*Sketch* platform (www.lessonsketch.org). For the items from each of the four obligations, see Appendices C through F. Each part of the questionnaire consisted of fifteen to eighteen scenarios that featured an undergraduate college mathematics classroom where an instructor chose to act on behalf of one of the professional obligations. Participants were then shown a statement about the instructor's action (e.g. "The teacher should stick to the mathematics at hand, rather than take class time to make connections to other mathematical ideas.") that they respond to on a 6-point Likert-type scale from strongly disagree to strongly agree. Each close-response question was followed by an open-response question asking participants to comment on their rating.

A version of this instrument has been used at the high school grade level with a national sample of mathematics teachers (Herbst & Ko, 2018; see also Boileau, Ko, & Herbst, 2018). The issue at play in every item remained the same, but some of the contextual cues were adjusted to be more appropriate for a college-level mathematics classroom (e.g., the instructor would say the topic needed to be finished by the end of the semester rather than the end of the school year or the content on the board would be calculus instead of geometry). Previously validated surveys still require additional reliability and validity evidence when administered in a different context (Rickards, Magee, & Artino, 2012). I expected the obligations and issues in the items remain relevant for college professors based on piloting. A sample of the new college-level scenarios were piloted in semi-structured interviews with 16 university mathematics instructors to ensure that the scenarios were realistic and plausible, and participants explicitly noticed what the items were designed to test. Two faculty and three graduate students that do research in mathematics education and had mathematics teaching experience at the undergraduate level, reviewed and critiqued all scenarios. A set of 20 linking items were administered to be able to compare responses from the high school instructor sample to the college instructor sample (see Vale, 1986).

Although college mathematics professors may differ from high school mathematics teachers in terms of how much they recognize each obligation, I expected the scales created for college professors would perform with similar reliability using the factor structures used for high school teachers. Each obligation is identified with a stakeholder, which all exist in the high school or college context. For each obligation, disciplinary, interpersonal, individual, and institutional, the mathematics, the society and classroom environment, the students, and the mathematics department or school/university are the respective stakeholders at either the high school or the college level. The instructors are all still held accountable to those four things, even if they interpret them differently.

Beliefs Instrument

To examine the role of beliefs, an instrument was needed that could measure beliefs directly linked to inquiry-oriented instruction. I chose a beliefs instrument by considering existing instruments that tested instructor beliefs with some relation to ideas about inquiry-oriented instruction. I considered three instruments that fit this description: (1) Wilkins (2008) asked about instructors' beliefs in predicting use of inquiry-based learning, (2) Stipek, Givvin, Salmon and MacGyvers (2001) used a beliefs instrument that included items on how teachers believe students learn mathematics (whether they focus more on correctness or students' understanding of concepts), and (3) Clark and colleagues' (2014) created items based on a broad sampling from theories about teaching and learning mathematics. The instrument used by Wilkins (2008) had the most direct link to inquiry-oriented instruction in that every item was directly asking if instructors believed that inquiry-based learning or aspects of inquiry-based learning were effective. However, what counted as inquiry-based learning was often unclear or not applicable to the college context such as asking about the effectiveness of using calculators, using computers, use of portfolios, and having students participate in hands-on activities. These questions seemed relevant to the elementary context but not to the undergraduate classroom context. The items from Stipek, Givvin, Salmon and MacGyvers (2001) seemed more relevant to the college context, but most of the constructs were not as clearly aligned with the use of inquiry-oriented practices. For example, some of the constructs in that instrument included an entity versus incremental view of intellectual ability, confidence in teaching mathematics, teacher control versus child autonomy, and extrinsic versus intrinsic motivation. I could imagine associations between these constructs and inquiry-oriented practices, but I did not see how these questions could help me answer whether or not beliefs aligned with practice. Instead, I would have had to ask something more akin to, "What beliefs align with practice?" This might have been valuable, yet an entirely different study. The instrument from Clark and colleagues (2014) seemed more appropriate for the college context and clearly linked to inquiry-oriented practices than the other two.

Clark and colleagues' (2014) sampling from theories on beliefs of teaching and learning led to constructs that I deemed would be reliable indicators of inquiry-oriented practices. The belief that students should be allowed to struggle aligns with learner-focused beliefs that view mathematics as an active constructive process, and content-focused beliefs that focus on developing students understanding of the content (Kuhs & Ball, 1986). Modeling for incremental mastery seems akin to content-focused beliefs that emphasize improving students' ability to carry out procedures (Kuhs & Ball, 1986). I suspected that the belief that students should struggle would predict more frequent use of inquiry practices, while belief that teachers should model would predict less. Instructors that believe they are aware of their students' dispositions could have more learner-focused beliefs, which again logically would predict more inquiry practices. Or, conversely, instructors that interact with their students more through things like group work, or hinting without telling, might perceive themselves as more aware of their students' dispositions.

For this study, I used all published beliefs items from Clark and colleagues (2018) (see Appendix B). The items were on a six-point scale from 1-Strongly disagree to 6-Strongly agree. Their study included an exploratory factor analysis that yielded three main factors consisting of 21 items grouping into the following three categories: belief that Teachers should Allow for

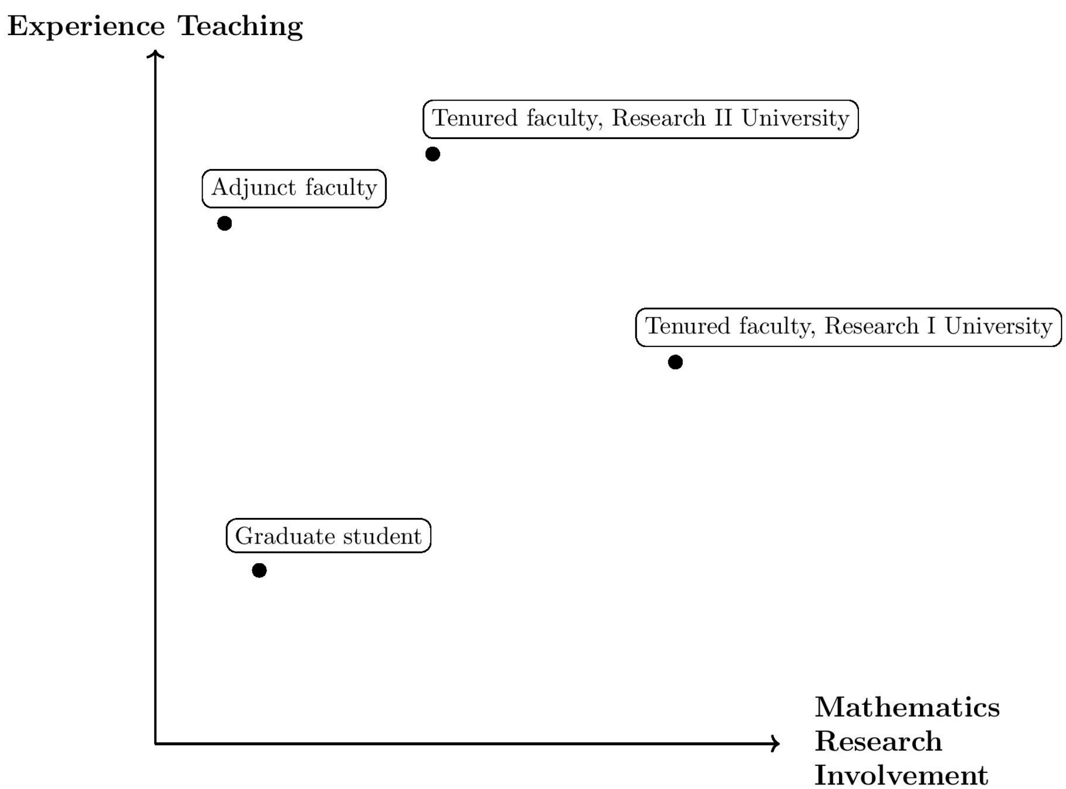
Student Struggle with Problems (TASSP), belief that Teachers should Model for Incremental Mastery (TMIM), and Teacher Awareness of Students' Mathematical Dispositions (TASMD). Their items had been tested with 259 upper-elementary and 184 middle-grade mathematics teachers. To keep the text of my results as transparent as possible, I use the names *Struggle, Model, and Awareness,* in place of the acronyms TASSP, TMIM, and TASMD, respectively. Though college mathematics instructors may have different beliefs than K-12 teachers, I expected the general factor structure should still hold.

Finally, all participants took a background information instrument to capture general characteristics of the sample. It included basic information such as years of teaching experience, status in the department, general research interests, and type of institution (see Appendix G). These variables are useful in analysis as possible controls or independent variables. The background variables were useful as a check if the sample was nationally representative, as sample design was based on voluntary participation.

Sample Design

I collected data from a national sample with participants representing 94 different mathematics departments across 37 states, plus Washington, D.C. I compiled a comprehensive list of Research I and Research II mathematics department emails based from an original list acquired by Pablo Mejía-Ramos of Rutgers University and updated by the GRIP lab, along with a handful of high-ranking liberal arts colleges. I asked the mathematics department secretary or chair to forward on a qualifying survey to anyone that instructed their courses. Participants qualified if they had at minimum one year of teaching experience and included tenure or tenure-track faculty, adjunct faculty, postdoctoral fellows, and graduate student instructors.

Figure 4.1: Hypothetical examples of sampling for variation in experience and research involvement



The two main covariates that I intended to capture by sampling a large variety of instructors were experience level with teaching and level of involvement in mathematics research, as demonstrated by the hypothetical profiles in Figure 4.1. All instruments were administered through Qualtrics and taken at the instructors' convenience. Participants were sent a $50 Visa gift card upon completion of the study. Once I collected the data in Qualtrics, I exported and merged it in Stata, and began analyses.

Background Characteristics

The background survey yielded the statistics for the participants in the study shown in Table 4.2 and Table 4.3. The table includes statistics for the total sample, that is, all participants of this study that completed at least a portion of the instruments. They did not necessarily complete all the instruments, but their data is used somehow in the results. This sample is representative of the national population of university mathematics instructors in terms of racial/ethnic groups. The Conference Board of Mathematical Sciences survey found that among all tenured, tenure-eligible, postdoctoral and other full-time faculty in mathematics departments of four-year colleges and universities, estimated racial/ethnic groups included 15% Asian, 3% Black, not Hispanic, 3% Mexican American/Puerto Rican/other Hispanic, 77% White, not Hispanic, 0% American Indian or Alaskan Native and native Hawaiian or Other Pacific Islander, and 3% unknown (Blair, Kirkman, & Maxwell, 2018). For my sample, participants self-reported race/ethnicity as 16% Asian, 1% Black or African American, 3% Hispanic or Latinx, 71% White, 1% American Indian or Alaska Native and Native Hawaiian or Other Pacific Islander, and 5% were other or preferred not to answer.

In contrast, the background survey for this study revealed that this sample was not representative in terms of gender. The sample used in this study includes more women than the national population of mathematics department instructors. In the annual survey of mathematical sciences in the US, 17% of the combined tenure and tenure-track faculty were female, and 33% of the doctorate-holding non-tenure-track faculty (including postdocs) were female (Golbeck, Barr, & Rose, 2018) in university mathematics departments. In this sample, 15/38 = 39% of combined tenure and tenure-track faculty are female, and 50/90 = 55.6% of non-tenure-track faculty are female (though this includes non-doctorate-holding faculty).

Table 4.2: Basic background statistics for study participants

	Percentage (%)	**Frequency (*n*)**
Years of experience teaching		
1 to 3	31.67	89
4 to 6	23.13	65
7 to 9	12.1	34
10+	33.1	93
Gender		
Female	49.81	133
Male	47.57	127
Another identity	1.12	3
Prefer not to answer	1.5	4
Position		
Graduate student	45.56	118
Postdoctoral researcher	8.88	23
Part-time faculty	5.02	13
Non-tenure-track faculty	25.87	67
Tenure-track faculty	7.34	19
Tenured faculty	7.34	19
None of the above	0	0

Table 4.3: Background characteristics of participants related to research level and teaching characteristics

	Percentage (%)	**Frequency (*n*)**
Carnegie classification		
Research I	66.29	175
Research II	21.59	57
Research III	1.89	5
Unsure	10.23	27
Research to teaching ratio		
0-25% research, 75-100% teaching	43.13	113
25-50% research	17.94	47
50-75% research	24.43	64
75-100% research	14.5	38
Heard of IOI or IBL		
Yes	56.55	151
Yes, but I don't know anything about it	24.34	65
No	19.1	51
Have used IBL		
Yes	41.22	61
No	44.59	66
Unsure	14.19	21
Enrollment for smallest course usually taught		
<25	38.64	102
26-45	46.59	123
46-100	11.74	31
100+	3.03	8

A final disparity to note is that this sample includes a greater proportion of graduate students than exists in the true national sample of undergraduate mathematics instructors. Graduate teaching assistants teach 4% to 25% of mainstream calculus courses in four-year colleges and universities in the US (Blair et al., 2018), yet they compose 45.56% of this sample.

Data Analysis

I confirmed the reliability of the scales in the instruments and explored the relationships between constructs using structural equation modeling (SEM). SEM refers to a family of related techniques, including factor analysis and structural regression modeling (Kline, 2016). It is a causal inference method that takes qualitative hypotheses and questions about causal relationships and outputs numerical estimates of hypothesized effects (Kline, 2016).

An essential step to establishing the reliability of a scale is to test the internal structure or dimensionality of instrument items (Furr & Bacharach, 2013). I used a combination of the software MPlus and Stata to conduct this analysis. Confirmatory factor analysis (CFA) is typically most appropriate when a theory-based hypothesis already exists or a test's dimensionality has already been established in a previous study, while exploratory factor analysis (EFA) is appropriate when developers are still working to understand the constructs at stake (Furr & Bacharach, 2013). I first tested the hypothesized internal structure of each instrument by conducting a confirmatory factor analysis for each instrument. Based on the theory and literature reviewed and the design of the instruments, I expected the PROSE items would be grouped by professional obligation, and the beliefs items would take on a three-factor structure grouped by allowing students to struggle, modeling for incremental mastery, and awareness of students' disposition beliefs. I expected the INQUIRE instrument would have a multidimensional structure – that is, it would have the constructs as outlined in Table 3.1, and those constructs would be correlated as part of a higher-order general factor of instructing with inquiry-oriented methods.

For any of the models that did not yield satisfactory fit statistics based on their hypothesized structures, I explored the structure of the items using an EFA and the information from the unsuccessful CFA. Once the dimensionality of each instrument was established, I checked the internal consistency of each factor by calculating Cronbach's alpha, mean inter-item correlations, and rho statistics. Though alpha scores are the most widely used metric of internal consistency (Furr & Bacharach, 2013), higher values can reflect a large number of items or excessive duplication of items rather than the desired quality of better internal consistency (Streiner, 2003). Therefore, I supplemented those scores with mean inter-item correlations and factor rho coefficients (Kline, 2016; Raykov, 2004). The factor rho coefficient "is the ratio of explained variance over total variance" (Kline, 2016, p. 313), which accounts for the varied contributions of each item loading. The work of establishing the reliability of each scale is a prerequisite to analyzing what inquiry-oriented practices are being practiced and what factors predict their use.

After establishing the scales that composed each construct, I explored what inquiry-oriented practices and patterns of practices existed through means of descriptive statistics and cluster analysis. Hierarchical cluster modeling is a method which, most often, starts with a specified number of groups *N*, and generates *N-1* groups by combining the two groups that result in the least increase in the within-group variation (Everitt, Landau, Leese, & Stahl, 2011). There exist various methods to reduce the within-group variation (Everitt et al., 2011). I used Ward's (1968) method due to its superior performance compared to other hierarchical clustering methods and widespread use (Mojena, 1977), which minimizes within-group variation by minimizing the error sum of squares. In hierarchical cluster modeling, the number of groups is commonly decided by looking at a dendrogram, a graph that shows the nested sequence of clustering (Everitt et al., 2011). The viewer decides a place to cut the dendrogram, eliminating the divisions of clusters that are below a certain height.

I used SEM to test what factors predict each inquiry-oriented practice. Each model included an inquiry-oriented practice as the dependent variable, and all three belief constructs and four professional obligations as independent variables. For the independent variables, I used parcels, that is, aggregated groups of items, rather than all the items themselves.

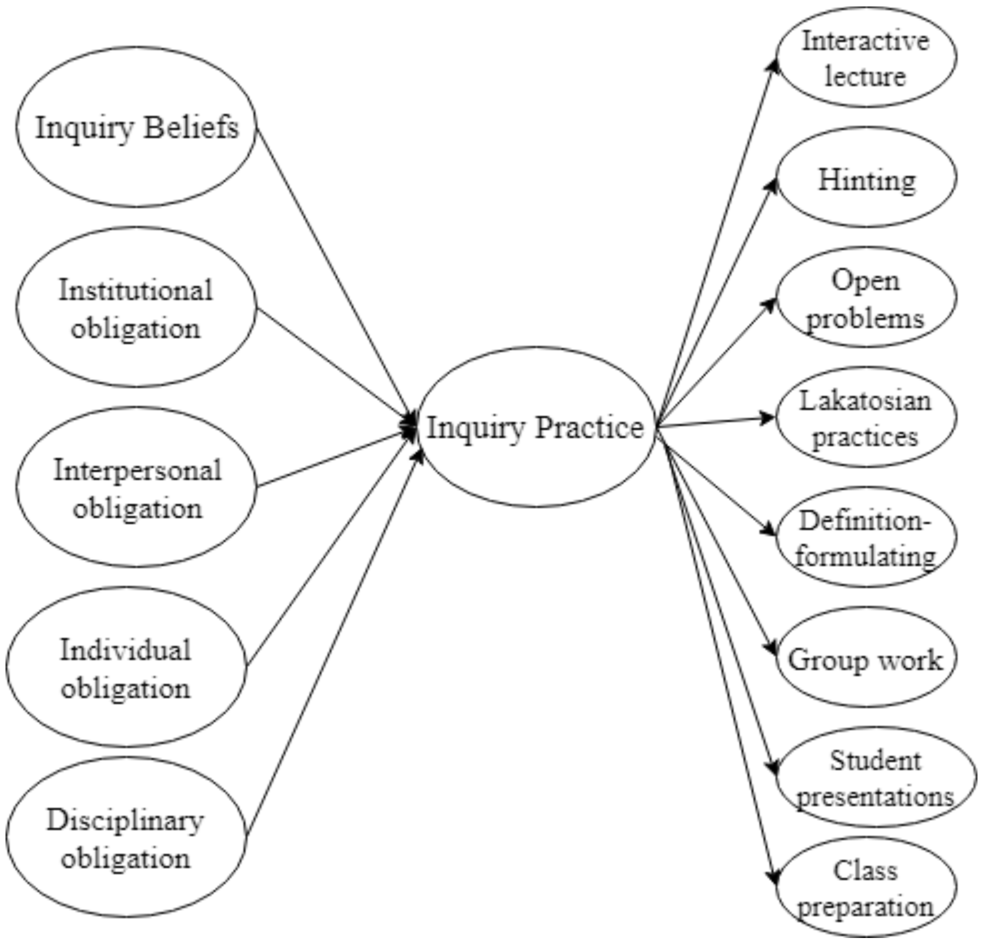
Following recommendation by Little, Cunningham, Shahar, and Widaman (2002), I used parcels only after exploring the dimensionality of the items to be parceled. Using parcels as indicators for latent constructs has the advantages of being more likely to fulfill the distribution assumptions inherent in SEM (that the data is distributed normally), and creating a model with fewer free parameters so that the sample size is sufficient (Little et al., 2002). Matsunaga (2008) recommended the use of three parcels per factor to keep the number of parcels minimal for the model fit, while still using enough parcels to avoid estimation bias. I used the simplest technique of randomly assigning items to these three parcels. These parcels indicated the latent constructs of beliefs, professional obligations, and inquiry-oriented practices that were interrelated in the following hypothesized conceptual model.

I used multinomial logistic regression to model what factors predict patterns of inquiry-oriented practice. The cluster analysis assigned each individual instructor to a group, which became the dependent variable. This type of regression was necessary because the dependent variable was nominal, not ordinal. The preliminary hypotheses about what the SEMs and multinomial logistic regression would look like were straightforward and brief.

Conceptual Model

The skeleton of the mediation model I originally wanted to test had beliefs predicting inquiry-oriented practice, with professional obligations acting as components that contribute to the variance in practice, as shown in Figure 4.2.

Figure 4.2: The conceptual model for the latent constructs measured by INQUIRE, PROSE, and STBIBT.



In practice, I found that IOI was not unidimensional, and that there were different patterns of usage for the different dimensions. It was more interesting to test the interactions from the latent constructs on the left to the constructs on the right, and patterns within the practices on the right. For example, I tested the relationship between beliefs and professional obligations directly on perceived frequency of using interactive lecture. Keeping the components of IOI separate allowed for the nuances and variation from practice to practice to not get lost, as they would have if they were compiled together into one score. The gist from the conceptual model that was valuable as a guide for analysis was the idea that beliefs and professional obligations are being tested as indicators of inquiry-oriented practice, which is composed of many distinctive components.

Chapter 5 **Scale Development**

For each instrument, I report efforts to construct and revise the scales used in subsequent sections. Item level descriptive statistics for each instrument can be found in Appendix H. For each instrument, I first examined the scale at an item level in Stata by viewing a correlation matrix and the item-rest correlations of each item. I used the correlation matrix to get a visual representation of what items were not grouping with their intended factors by highlighting all correlations higher than a certain number. The item-rest correlations gave the correlation between scores of a given item and the aggregated score of the remaining items – thus quantifying how well the item grouped with the others. If there were any items that were extreme outliers from their respective groups, I checked if there was a theoretical reason and, if so, removed the item before continuing.

If the scale had been developed previously or if there was a strong theoretical grounding, I tried a CFA using the hypothesized structure. I conducted the analysis in MPlus using the WLSMV estimator, which does not assume normally distributed variables, because the data are categorical and ordinal (Brown, 2006). All factor loadings are reported in the appendices in STDYX standardization, so that a maximum value for a factor loading of 1 would indicate that an item contributes perfectly (without error) to the existence of a latent construct. I removed items with low factor loadings or redundancy indicated by high correlations or modification indices. I relied on commonly used global fit indices as guidelines: the Root Mean Square Error of Approximation (RMSEA)$\leq .05$, the Comparative Fit Index (CFI)≥ 0.95, and the Tucker Lewis Index (TLI)≥ 0.95 (Hu and Bentler, 1999). If two similar, non-nested models seemed conceptually possible, I used the Bayesian information criterion (BIC) to choose the preferred model (Raftery, 1995).

In one instance, namely for the student-content constructs from the INQUIRE instrument, I used a split-sample EFA (Kline, 2016), because I suspected a slightly different item grouping based on patterns in the correlation matrix. The EFA was conducted on a randomly selected half of the sample. Once a model was determined, it was confirmed on the second half of the sample to ensure the structure was not a product of overfitting the data.

To ascertain the dimensionality of the model, I used a combination of the theory behind the design of the instrument and eigenvalues, which indicate the amount of variance explained by a given number of factors. The more factors in the model, the more variance is explained – but with each additional factor there are diminishing returns. That is, one can imagine the extreme case where there are as many factors as variables, in which case all variance is explained but the reliability gained by using multiple items to assess a construct is lost. I used a scree plot[3] and Kaiser's criterion (Kaiser; 1960, Cattell, 1978) to decide the minimum number of factors to create meaningful latent constructs and still account for enough variance. Kaiser's criterion (1960) recommends an eigenvalue cutoff of 1.0, while Cattell (1978) recommends to look for an "elbow" in the scree plot, to see when adding additional factors stops generating substantial gain. I looked at oblique factor loadings and considered if there were conceptual reasons for the groupings that emerged both to help determine the number of factors and to make sense of how the items are functioning for making improvements in future iterations of the instruments.

[3] A scree plot is a graph of eigenvalues. The possible number of factors lies on the *x*-axis, and the *y*-axis consists of corresponding *y*-values (Zhu & Ghodsi, 2006).

After determining the items that compose each factor (of unidimensional and multidimensional models), I checked the internal consistency of each of the scales. Commonly used measures of internal consistency are Cronbach's alpha, mean inter-item correlations (IICS), and the factor rho coefficient (Kline, 2016). A common guideline for Cronbach's alpha is to consider values over .7 as acceptable and those below .5 as unacceptable (Kline, 2016), for IICs is to consider values between .15 and .50 as acceptable (Clark & Watson, 1995), and for rhos to be above .7 (Hair et al., 2006).

Missing Data

The INQUIRE instrument had a flawed implementation: In Qualtrics, if a user responded to the question, "How often do you ask students to revise a definition?" with "1-Never", they were skipped past the remaining sections of the survey (all student-student and teacher-content questions, and all questions on upper-division courses) due to an error in the survey logic. There were initially 252 finishers of the INQUIRE instrument. Of these, 143 chose "1-Never". I have complete data for 194 finishers, that is, I recovered data for 85 participants, who retook the rest of the instrument or who took the instrument after the error had been detected and corrected. Fifty-eight participants did not respond to solicitations to finish the instrument. I call the sample including the 194 finishers and the 58 participants who did not finish the *total sample*. I call the sample with the 194 finishers but without the 58 participants who did not finish *partial sample A*. I call the sample without those that chose 1-Never, of 252-143=109 participants, *partial sample B.* Note that the increased attrition from the lower-division questions to the upper-division questions is not solely due to this error, but largely because many instructors (115) had never taught an upper-division course.

To assess the impact of the missing data, for analyses that have complete data for everyone including the unfinished participants, I present side-by-side results with the effective sample and partial sample A (dropping participants that did not fully complete the INQUIRE instrument). For analyses that involve the INQUIRE items without complete data for the 58 participants, I present side-by-side results for the effective sample for which data exists and partial sample B which excludes those that chose 1-Never. This comparison is a transparent way to display any potential impacts and biases of this error.

INQUIRE

The INQUIRE instrument was taken twice by participants, once pertaining to a lower-division course and once for an upper-division course, and contained nine hypothesized latent constructs. For the sake of the scale development, I focus on lower-division responses because there were fewer participants who had taught upper-division courses so the sample size is not sufficiently large. In addition to the errors in implementation that resulted in missing data, I also had an error with scales. I initially designed the survey to have 5-point Likert responses instead of 6, and when I changed them I missed some questions. In the descriptive statistics and reliability scores, I rescaled the 5-point questions to 6-point so that scores can be compared. For scale development and factor analyses, I did not rescale the scores because they are all standardized.

I first detail the development of each grouping of constructs by their organization in the instructional triangle. Then I do a factor analysis using all the INQUIRE constructs to ensure that the constructs across the full instrument still function as separate constructs as conceptualized when placed all in the same model.

INQUIRE: Student-Content practices

For the items relevant to student-content practices, I expected items to form the following four factors: *open problems, constructing, critiquing,* and *definition-formulating*. A CFA using the hypothesized structure did not yield ideal fit statistics (RMSEA=.09, CFI=.93, TLI=.93). After removing two redundant items, the fit statistics were acceptable (RMSEA=.08, CFI=.95, TLI=.95). However, the correlation matrix reflected that items about definition-formulating grouped with some of the constructing and critiquing items and factors still had high correlations (constructing with definition-formulating: .91; critiquing with definition-formulating: .90, critiquing with constructing: .90). Conceptually, I could imagine different groupings for the items – for example, all items about proofs that were originally placed in different groups (in definition-formulating, critiquing, and constructing) had high correlations. Thus, I tried an EFA on a randomly chosen half of the sample (n=128). Eigenvalues for the first five factors were 10.977, 1.847, 1.568, 1.199, and 1.065, respectively.

I chose to use a 3-factor model due to the diminishing returns shown by eigenvalues and the theoretical interpretation of three factors. The four-factor model did not have such a clear interpretation. The three latent constructs that emerged included the instructor providing opportunities for students (1) to *construct or critique* conjectures and claims (2) to *prove* claims via proofs and previously established results, and (3) to *solve open problems*, that is, problems that do not have only one solution or one solution method. The fit indices for the EFA were acceptable (RMSEA=.08, CFI=.96, TLI=.94). The EFA also suggested that the model would fit better with the removal of two items that were redundant (see Appendix A).

Table 5.1: Comparison of factor loadings for student-content items in the INQUIRE instrument with the total sample and partial sample A

Item	Loadings for total sample, *n=252*	Loadings for partial sample A, *n=194*	Difference
Construct/Critique			
LSC_1_DEFINE	0.74	0.74	0
LSC_2_DEFINE	0.81	0.81	0
LSC_3_DEFINE	0.85	0.85	0
LSC_1_LAKATOS	0.78	0.80	0.02
LSC_2_LAKATOS	0.82	0.82	0
LSC_3_LAKATOS	0.78	0.79	0.01
LSC_4_LAKATOS	0.61	0.64	0.03
LSC_8_LAKATOS	0.74	0.77	0.03
LSC_10_LAKATOS	0.78	0.80	0.02
LSC_11_LAKATOS	0.69	0.72	0.03
Prove			
LSC_4_DEFINE	0.65	0.64	-0.01
LSC_5_LAKATOS	0.70	0.72	0.02
LSC_6_LAKATOS	0.71	0.74	0.03
LSC_7_LAKATOS	0.69	0.67	-0.02
LSC_9_LAKATOS	0.67	0.64	-0.03
LSC_12_LAKATOS	0.69	0.67	-0.02
LSC_13_LAKATOS	0.84	0.86	0.02
LSC_14_LAKATOS	0.85	0.85	0
Open Problems			
LSC_1_OPEN	0.60	0.63	0.03
LSC_3_OPEN	0.57	0.57	0
LSC_4_OPEN	0.59	0.62	0.03
LSC_5_OPEN	0.77	0.75	-0.02

This factor structure was tested with a CFA on the other half of the sample (*n=124)*. The model fit reasonably well (RMSEA=.07, CFI=.96, TLI=.96) with significant loadings ranging .60~.92. The correlation between factors was high, which might be due to the similar nature of the mathematical activities. For instance, constructing and critiquing is somewhat similar to the activity of putting together a proof. Finally, I ran the model with the total sample and the incomplete sample to compare factor loadings and gain insight into any potential impact of the missing data, shown in Table 5.1.

INQUIRE: Teacher-Student Practices

For the constructs related to teacher-student interactions, I expected items to form the following two factors: *interactive lecture* and *hinting without telling.* The fit statistics for the theorized structure were not good (RMSEA=.12, CFI=.86, TLI=.83) but the items seemed to group into the hypothesized two factors. I removed items with high covariance in order to avoid redundancy in the scale (see Appendix A). The removed items were all designed to measure interactive lecture, and asked similarly worded questions about how often instructors asked students to work on problems during their lecture. For example, the items "While teaching the whole class, how often, after demonstrating how to solve a problem, do you ask students to try a similar problem?" and "While teaching the whole class, how often do you pause your presentation to ask students to work on a problem or problems?" were redundant. The final model had acceptable fit (RMSEA=.06, CFI=.97, TLI=.96) with loadings ranging .41~.89. Comparisons between the loadings for partial sample A and the total sample are shown in Table 5.6.

Table 5.2: Comparison of factor loadings for student-teacher items in the INQUIRE instrument with the total sample and partial sample A

Item	Loadings for total sample, *n*=262	Loadings for partial sample A, *n*=194	Difference
Interactive Lecture			
LTS_2_LECTURE	0.54	0.50	-0.04
LTS_3_LECTURE	0.65	0.70	0.05
LTS_4_LECTURE	0.52	0.47	-0.05
LTS_5_LECTURE	0.49	0.47	-0.02
LTS_6_LECTURE	0.57	0.51	-0.06
Hinting without telling			
LTS_1_HINT	0.71	0.66	-0.05
LTS_2_HINT	0.89	0.95	0.06
LTS_4_HINT	0.42	0.43	0.01

INQUIRE: Student-Student Practices

The constructs related to student-student interactions grouped into their hypothesized factors, student presentations and group work, with acceptable fit (RMSEA=0.14, CFI=0.94, TLI=0.93). The correlation between factors was high (0.707), indicating that there might be one underlying factor. However, the fit of the one-factor model (BIC=7729.05) in comparison to the two-factor model (BIC=7440.56) demonstrated that the two-factor model was preferable. I removed item lsspresent_4 (How often do you have students take leadership roles during class?), because it is the only item that conceptually could load onto both factors, as leadership roles could occur within whole group or small group discussions. I also removed lssgroup_2 and lssgroup_6 because they were redundant with other items and cross loaded on the student presentation construct.

Table 5.3: Comparison of factor loadings between the total sample and partial sample B

Item	Loadings for total sample*, *n*=194	Loadings for partial sample B, *n*=107	Difference
Presentations			
LSS_1_PRESENT	0.85	0.85	0
LSS_2_PRESENT	0.88	0.84	-0.04
LSS_3_PRESENT	0.92	0.91	-0.01
LSS_5_PRESENT	0.85	0.87	0.02
LSS_6_PRESENT	0.80	0.75	-0.05
LSS_7_PRESENT	0.76	0.70	-0.06
Group work			
LSS_1_GROUP	0.68	0.75	0.07
LSS_3_GROUP	0.87	0.89	0.02
LSS_4_GROUP	0.86	0.92	0.06
LSS_5_GROUP	0.85	0.85	0
LSS_7_GROUP	0.79	0.83	0.04
LSS_8_GROUP	0.79	0.78	-0.01
LSS_9_GROUP	0.81	0.80	0

*Note that for the student-student responses, the total sample is the same as Partial Sample A

The final model fit well (RMSEA=.10, CFI=.97, TLI=.97) with loadings ranging .68~.92. Factor loadings in Table 5.3 are shown for the total sample (including all the data that I collected), and for partial sample A set that removes the participants who retook the instrument after the logic error was discovered. The comparison demonstrates that differences between the results with or without those who claimed to never have students revise definitions are trivial.

INQUIRE: Teacher-Content Practice

The teacher preparation construct from the teacher-content interaction category did not fit well into a unidimensional model (RMSEA=.28, CFI=.88, TLI=.83). I removed ltcpresent_8 (“How often do you design your lesson to include experiences you have had learning mathematics?”) because it was highly correlated (.91) with ltcprsent_9 (How often do you design your lesson to include experiences you have had doing mathematics?”) .

Table 5.4: STDYX Loadings for all items in the final model for the INQUIRE instrument, n=*257.*

Construct	Item	Loading	SE
Constructing/Critiquing	LSC_1_DEFINE	0.73	0.04
	LSC_2_DEFINE	0.79	0.03
	LSC_3_DEFINE	0.85	0.03
	LSC_1_LAKATOS	0.78	0.03
	LSC_2_LAKATOS	0.81	0.03
	LSC_3_LAKATOS	0.76	0.03
	LSC_4_LAKATOS	0.62	0.04
	LSC_8_LAKATOS	0.76	0.03
	LSC_10_LAKATOS	0.8	0.03
	LSC_11_LAKATOS	0.71	0.04
Proving	LSC_4_DEFINE	0.67	0.04
	LSC_5_LAKATOS	0.68	0.04
	LSC_6_LAKATOS	0.75	0.04
	LSC_7_LAKATOS	0.68	0.04
	LSC_9_LAKATOS	0.66	0.04
	LSC_12_LAKATOS	0.69	0.04
	LSC_13_LAKATOS	0.84	0.03
	LSC_14_LAKATOS	0.84	0.03
Open problems	LSC_1_OPEN	0.56	0.05
	LSC_3_OPEN	0.56	0.05
	LSC_4_OPEN	0.64	0.05
	LSC_5_OPEN	0.77	0.04
Interactive lecture	LTS_2_LECTURE	0.5	0.07
	LTS_3_LECTURE	0.66	0.07
	LTS_4_LECTURE	0.51	0.08
	LTS_5_LECTURE	0.64	0.08
	LTS_6_LECTURE	0.27	0.09
Hinting without telling	LTS_1_HINT	0.55	0.07
	LTS_2_HINT	1.03	0.09
	LTS_4_HINT	0.48	0.07
Student presentations	LSS_1_PRESENT	0.82	0.04
	LSS_2_PRESENT	0.88	0.03
	LSS_3_PRESENT	0.9	0.03
	LSS_5_PRESENT	0.82	0.03
	LSS_6_PRESENT	0.8	0.05
	LSS_7_PRESENT	0.88	0.04
Group work	LSS_1_GROUP	0.63	0.05
	LSS_3_GROUP	0.86	0.02
	LSS_4_GROUP	0.84	0.03
	LSS_5_GROUP	0.86	0.03
	LSS_7_GROUP	0.78	0.03
	LSS_8_GROUP	0.82	0.04
	LSS_9_GROUP	0.85	0.03

The items were almost identical, designed with the idea that learning and doing mathematics might mean different things. Even with this item removed, the model did not fit well.

Using an exploratory factor analysis, eigenvalues indicated that the first three factors explained enough of the variance (1 factor: 3.83, 2 factors: 1.56, 3 factors: 1.21, 4 factors: .81). Solutions for two, three, and four factors were each analyzed. With nine items, a three-dimensional model does not maintain enough items per factor to create reliable latent factors although the model had acceptable fit (RMSEA=.17, CFI=.98, TLI=.94). The characteristics of the items that load onto each factor (see Table A.5) can help inform how to develop the constructs under this relationship for future iterations of the INQUIRE instrument. The first and third factor are composed to two items each. The first factor has two items that ask about preparing worksheets. The third factor is focused on the two previously mentioned redundant items (ltcpresent_8 and ltcpresent_9) that dealt with preparing a lesson to include experiences the instructors had with mathematics. Four items loaded onto the second factor, all dealing with design and/or searching for appropriate problems. Future versions of the instrument could include more items on each of these three constructs, or merge the worksheet and problem items by using a word that encompasses both problems and worksheets, such as materials. Due to the lack of items and insufficient theoretical backing for the factors that emerged, I did not incorporate the teacher-content latent constructs in the remainder of my analyses.

INQUIRE: Entire Model

The model including all seven INQUIRE constructs from the three relationships of the instructional triangle without any revisions from the previous models (e.g., no added cross-loadings or correlations) had acceptable fit (RMSEA=.05, CFI=.94, TLI=.94). All loadings and standard errors from that model are shown in **Error! Reference source not found.**. The final structure of constructs measured by the INQUIRE instrument after psychometric work is shown in Table 5.5.

Table 5.5: The final list of constructs measured by the INQUIRE instrument after the psychometric work.

Triangle Relationship	Constructs	Description
Student-Content	Open problems	Posing problems that either have multiple solutions or multiple nontrivial ways of arriving at a solution
	Constructing/Critiquing	Posing tasks that ask students to create or critique conjectures or claims
	Proving	Using previously established results to construct formal arguments for mathematical claims
Teacher-Student	Interactive lecture	Instructing to the full class while asking for feedback from students, asking questions of students, and having students engage with the mathematics
	Hinting without telling	Guiding a student to work productively without directly telling the student a correct way to proceed
Student-Student	Group work	Creating an environment where students work together on mathematical tasks or problems
	Student Presentations	Having a student or students present completed or in-progress work to the class

The constructs verified in this section give construct validity to the way IOI was conceptualized in the theoretical framework and methods chapters. The development of the next two instruments enable the ability to predict the constructs in the INQUIRE instrument in the following chapter.

Beliefs

The exploratory factors from Clark et al. (2014) beliefs instrument - belief that students should be allowed to struggle, belief that teachers should model for incremental mastery, and awareness of students' mathematical dispositions - did not fit well as a structure for the data collected (RMSEA=.10, CFI=.81, TLI=.79). Without looking at the data, I theorized a different factor structure based on the content of the items and consulting with colleagues. This theoretical model performed better (RMSEA=.09, CFI=.86, TLI=.85).

Table 5.6: Comparison of factor loadings of the total sample and Partial Sample A for the beliefs instrument

Item	Loading for total sample, *n*=243	Loading for Partial sample A, *n*=184	Difference
Struggle			
Q1_1	0.72	0.71	-0.01
Q1_2	0.74	0.71	-0.03
Q1_3	0.57	0.65	0.08
Q1_4	0.58	0.56	-0.02
Q1_6	0.61	0.66	0.05
Model			
Q1_5	0.58	0.50	-0.08
Q1_7	0.83	0.82	-0.01
Q1_8	0.54	0.54	0
Q1_9	0.7	0.68	-0.02
Q1_10	0.51	0.45	-0.06
Q1_12	0.58	0.65	0.07
Q1_13	0.37	0.42	0.05
Awareness			
Q1_14	0.78	0.75	-0.03
Q1_15	0.85	0.84	-0.01
Q1_16	0.84	0.83	-0.01
Q1_18	0.29	0.24	-0.05

I dropped two items with high cross-loadings and three that were redundant, with details recorded in Table A.6 in Appendix B. After dropping items, only one item, q1_5 ("To teach mathematics, first model the activity, then provide some practice and immediate feedback, and, finally, clarify what the assignment is and how it is to be completed."), did not load onto the original factor prescribed in Clark et al. (2014). It had been associated with struggle as an item with negative loading, but I thought that conceptually it was more appropriately grouped with the construct about modeling. Although the factor analysis did not yield the exact same model as

Clark et al. (2014) had proposed, the instrument has the same functionality. That is, a transposed item and a five dropped items did not change the integrity of the constructs being measured, as the indicators in the final model still captured how the original constructs were described. The final model fit well (RMSEA=.05, CFI=.96, TLI=.96).

PROSE Instruments

For each of the PROSE instruments, I first ran unidimensional confirmatory model. I wanted to create a scale that could be used consistently across other datasets to reflect each obligation while maximizing a variety of features of the obligation captured by each item, so I was conservative with the items that I dropped. There exist four other versions of the PROSE instruments, catering to different grade levels, so I avoided removing linking items as much as possible so that the scales could be compared against each other.

PROSE-Individual

The scenarios in the PROSE-individual obligations instrument featured an instructor breaching some norm on account of attending to individual students' needs. The recognition of the individual obligation is measured by whether the instructor agrees that the breach was appropriate or not. Before finalizing a unidimensional CFA for recognition of the individual obligation, I removed items that did not contribute meaningfully to the scale. I was able to check for the reasons behind low performance of items by examining open responses. I first removed item A4145, as participants had an almost universal objection to the scenario presented in the item. The scenario featured a student who did not want to follow instructions to share their solution, and participants responded to the statement, "The teacher should wait to speak with the student's advisor about the lack of cooperation rather than immediately lower the student's participation grade". The item was meant to elicit reactions to attend to the individual's needs by reaching out to the advisor. But instead, participants had a strong reaction to the instructor's behavior and the idea that an advisor would be involved or grades would be affected. Participants remarked that the instructor was being too aggressive, and were surprised that students have advisors or participation grades could even be altered. The study participants were attending to the individual student by worrying about the instructor's actions, but not in the way that was intended or in a consistent direction. That is, participants were attending to the individual obligation by both agreeing with the statement (e.g., by disliking the disruption of the students' autonomy by going to an advisor) and disagreeing with it (e.g., by feeling like a student should not be subject to losing participant points for this). These inconsistent reactions were reflected with the lowest item-rest correlation (.15). After removing this item, the fit statistics of the unidimensional model were still poor (RMSEA=.09, CFI=.68, TLI=.64). I recommend that future iterations of this item should use something different from an advisor that is more appropriate in a college context – such as waiting to speak with the student one-on-one later in office hours.

I removed item A4035 because it did not all elicit responses on the desired issue, which was whether or not to attend to the needs of an individual student by giving them extra time. The item featured a class who was about to take a quiz. A student who had been talking and not paying attention asked for more time to copy down a review problem on the board before beginning the quiz. Many responses indicated that their decision depended on things that were not clear from viewing the scenario. For example, responses depended on whether students had to put away their notes. If they were supposed to have put away their notes, a student could get it from another student later. But if they were not, there was no reason the student could not keep copying it down while other students began the quiz. Other participants objected to the instructor teaching content so close to the quiz and handing out the quiz while students were still taking notes. It was unclear how this confusion and objections would have translated into their closed-response choice. Though there were some open responses that were relevant to recognizing the individual obligation, there was also a lot of noise which led to this item having a low item-rest correlation (<.3). To remove some of the item's noise for future iterations, I would have the teacher explicitly ask students to put away their notes and also wait to hand out the quiz until the students had copied the material down as part of the given scenario.

I removed items A4055 and A4125 because they had such a small variance that they did not contribute to the scale. Item A4055 was intended to see if the instructor recognized the individual obligation by taking extra time to answer a students' tangentially related question. The instructor gives a story problem that includes the word meniscus. Participants generally felt that the instructor should obviously explain what it is to the student who asked. With few exceptions, all participants recognized the individual obligation in this item, with the highest mean response (M=5.11) of all the individual obligation items, and had a low item-rest correlation (<.3). A possible way to make the item more difficult would be to make the source of confusion more complicated to explain. With a meniscus, participants expected instructors to be able to define it in a couple of short sentences. With something more complicated and time consuming to explain, it might make responses to the dilemma more varied. In item A4125, a student is on their phone and the participants are asked to respond to the statement, "The teacher should complete the example without addressing the inattentive student rather than have the student move to the front of the room and restart the example." From reading through open responses, there seems to be an aversion among participants to address behavioral issues, and rather feeling like they should give the student autonomy to choose to not pay attention. Almost every participant strongly agreed with this statement. It had one of the lowest mean recognition scores (*M*=2.62) and a low item-rest correlation (<.3). A way the item could be improved is to make the question more difficult to explain – like something the teacher could not explain in less than a minute.

I removed the items A4025, and A4045 because they both had reactions from instructors that showed attention to aspects that were not relevant to the intended design of the items. Item A4025 was designed to target whether or not an instructor was worried about potentially embarrassing a student by bringing up their mistake with the whole class. However, instructors' open responses varied depending on whether they believed that the calculator error would be a common one or not. If they believed the error was uncommon, they said they wouldn't need to bring up the issue with the class, while if they believed it was common, they thought there would be value in addressing it with the class. They were attending to their knowledge of the error, rather than attending to the emotional well-being of the student who made the mistake. The item consequently had a low loading (<.3). For future iterations of the instrument, the item could be revised to make the error more obviously common, so that the variance would come from the potential embarrassment of the student rather than judgement about the error.

Item A4045 concerned a student who had finished early and wanted to work ahead. It was designed to have those that recognized the individual obligation agree to let the student work ahead at their own pace. However, many participants noticed issues having to do with group work, and how it could benefit the student to work with a partner. In that way, a large portion of participants were finding a way to recognize the individual obligation but while selecting a score that would reflect less recognition of the individual obligation. This could explain why the item had a low loading (<.3). Future versions of the item could indicate that a pair of students were ready to move ahead, so that agreeing with the item would be attending to the individual student's need to go at their own pace as well as have the benefits of working with a partner.

Table 5.7: Comparison of factor loadings for the total sample versus the partial sample A for the PROSE-individual instrument

Item	Loading for total sample, *n*=236	Loading for Partial sample A, *n*=173	Difference
A4165	0.39	0.47	0.08
A4115	0.29	0.29	0
A407L	0.27	0.22	-0.05
A4105	0.55	0.59	0.04
A402L	0.48	0.42	-0.06
A4085	0.60	0.63	0.03
A404L	0.52	0.50	-0.02
A4075	0.59	0.49	-0.1
A4095	0.49	0.57	0.08
A401L	0.29	0.32	0.03
A405L	0.71	0.66	-0.05
A403L5	0.60	0.52	-0.08

I removed the aforementioned items with problematic issues, achieving a final unidimensional model with acceptable fit statistics (RMSEA=.07, CFI=.93, TLI=.91) with significant loadings ranging .27~.71. Details including examples of two full scenarios, text for each item, and item loadings or details as to why they were dropped are recorded are shown in Appendix C. Loadings for the total sample are shown in Table 5.7. I include loadings from partial sample A (excluding the participants that did not finish all instruments) for comparison.

PROSE-Interpersonal

The scenarios in the PROSE-interpersonal obligations instrument featured an instructor breaching some instructional norms on account of improving the social environment of the classroom. The recognition of the interpersonal obligation is measured by whether the instructor agrees that the breach was appropriate or not. For the unidimensional CFA, I started by removing A3175 and A3185, as both did not function as designed. Item A3175 was intended to test if participants recognized the interpersonal obligation by stopping a student from sharpening their pencil, in the interest of protecting the shared aural environment for everyone in the class. In open responses, participants often thought the instructor should let the student sharpen their pencil in a way that showed they recognized the individual obligation towards that student, but that did not provide evidence of whether they cared or did not care about attending to the interpersonal environment. For example, participants said the student might need to take notes and not have another writing utensil, and that stopping the student could embarrass them. There were also comments on the unrealistic notion of a room still having a pencil sharpener. Possibly due to these comments, the item had a low item-rest correlation (.18) with the other items meant to probe recognition of the interpersonal obligation. An improved item could feature an aural disruption more realistic for a college classroom environment (e.g., noises from a phone or other electronic device) and see whether participants deem it appropriate to stop the student from using it.

Item A3185 was designed to see if instructors would recognize the interpersonal obligation by asking some disruptive students to stop talking among themselves during class, thereby attending to creating a classroom environment where everyone could listen. Most open responses expressed a desire for the instructor to have asked the chatting students if they had a question, and showed concern for the individual understanding or shyness of the talking students, rather than reprimanding them. The open responses did acknowledge a need to attend to the interpersonal environment, but in a way that would also include the needs of the chatting students. Thus, many of the participants disagreed with the item and did not recognize the interpersonal obligation, because they would have liked to attend to the interpersonal obligation in a different manner – not because they did not recognize the interpersonal obligation. This item had the lowest item-rest correlation (-.02) and I would improve it by having the student be whispering about something explicitly off-topic, so that the concern about students having questions would not obscure the choice to attend to the classroom environment.

Items A3095 and A3125 had response patterns distinct from the core group of interpersonal PROSE items. Responses to item A3095 did not vary enough, rather they showed a widespread rejection of the appropriateness of the instructor's actions. The item was intended to probe whether instructors would recognize the classroom environment by having a student, who asked to take a midterm early, check with the rest of the class to see if it was okay. Open responses reflected that the answer to taking it early should not depend on other students, that opening a discussion on something like an exam date is dangerous in terms of class management, and that the teacher (rather than the student) could check to see if anyone would like to join the student taking the exam earlier in order to protect the student's anonymity or personal reasons for needing to be absent. Absent from these reasonings were ideas about the importance of creating a classroom environment where opportunities like these are discussed and shared. The lack of variance overall to the item resulted in a low factor loading (<.3). An improved item could have a student check with the class to see if something less high stakes was okay with the rest of the class, such as the time of a practice exam or exam review session, turning off an extra bright light, or changing rooms so that the chalk board was more visible from all parts of the room. That way the instructors could decide to attend to the interpersonal environment without having the confounding issues of the danger or impossibility of discussing exam dates.

Item A3125 was designed to see if the instructor would attend to the well-being of the class by moving a physically strenuous activity from outside to inside the hallway on a hot day. Many responses questioned whether an activity like this would happen in a college-level course. The corresponding closed responses with the open responses of questioning did not show a consistent pattern (e.g., responses could range between 2 and 4 without a clear reason why). Responses also reflected that their choice depended on how hot it was outside and whether it would disrupt other classes if it took place in the hallway, creating extra variation in responses. Possibly for these reasons, the item did not load well (<.3). In future iterations, there might be a more believable version of an activity, and subsequent compromise to make it more inclusive, that would be more appropriate for the college setting. For example, a more realistic scenario could involve an activity running a simulation but some students not having the technology to do so, so somehow modifying the activity accordingly to include everyone.

Items A3085 and A301L were removed because they were redundant with other items. Item A301L was redundant with A3115 and A305L. They all had to do with attending to the interpersonal environment by calling on a new student, rather than one who was already sharing something. Item A3085 was designed to test whether instructors would recognize the interpersonal environment of the classroom by having some students relocate where they are working. It was redundant with A306L and A3075, which both were designed to see if instructors would recognize the interpersonal environment by attending to other students who were not as eager to participate. The content of these three items was not necessarily equivalent, but the response patterns were too similar for A3085 to give enough added information for the cost of the extra parameter. While it is useful in the development of a scale to include redundant items, the final version should take out ones that are excessively redundant (DeVellis, 2016). I recommend including these items in future uses of the instrument, but with the caveat that they may eventually need to be removed to create scores.

Table 5.8: Comparison of factor loadings for the total sample versus the partial sample A for the PROSE-interpersonal instrument

Item	Loading for total sample, *n*=205	Loading for partial sample A, *n*=160	Difference
A3015	0.43	0.47	0.04
A3115	0.44	0.38	-0.06
A304L	0.59	0.63	0.04
A3145	0.34	0.33	-0.01
A306L	0.54	0.55	0.01
A3075	0.37	0.38	0.01
A3055	0.65	0.62	-0.03
A3105	0.61	0.67	0.06
A305L	0.54	0.58	0.04
A3035	0.73	0.75	0.02
A3025	0.52	0.53	0.01
A3065	0.67	0.65	-0.02
A3045	0.64	0.56	-0.08
A3165	0.49	0.54	0.05
A302L	0.67	0.66	-0.01
A3135	0.47	0.43	-0.04
A303L	0.60	0.60	0
A3155	0.43	0.43	0

The final model retained 18 items with acceptable fit (RMSEA=.07, CFI=.93, TLI=.92) with significant loadings ranging .34~.73 (see Table 5.8). Details including examples of two full scenario, text for each item, and item loadings or details as to why they were dropped are recorded are shown in Appendix D. I ran this model with partial sample A (all participants except those that did not finish the instruments) as well and include loadings in the same table for comparison.

A final note about the PROSE-interpersonal items is that while the items do address issues around making the class an environment where everyone can learn, such as by distributing participation among students and making sure everyone in the class has a chance to understand material, they do not directly address issues of making the classroom more equitable for underrepresented groups. For example, there are no scenarios of the flavor of asking students to switch groups so that there would not be a lone female in a group. This is an area where the PROSE instrument could be expanded to cover what is meant by the interpersonal obligation more comprehensively.

PROSE-Disciplinary

The scenarios in the PROSE-disciplinary obligations instrument featured an instructor breaching some norm on account of representing mathematics more authentically. The recognition of the disciplinary obligation is measured by whether the instructor agrees that the breach was appropriate or not. I started by excluding item A1035 because it had little variance shown both in the open responses and item-level statistics. It was designed to see if instructors would recognize the disciplinary obligation by explaining the reason behind why $0! = 1$ instead of saying it was arbitrarily defined and moving on, which would show some effort to represent the field of mathematics accurately. Almost all instructors believed it was beneficial to explain it, saying in their open responses that students need to know that definitions have rationales and mathematics is not just jumble of arbitrary definitions. Due to the universal agreement, the item had a low item-rest correlation (-.03).

I also removed A1075 due to the instructors not responding as intended. The item had been created to see if instructors would correct a students' work on the board to be mathematically accurate, thereby recognizing the disciplinary obligation, or otherwise continue on with the problem without making the correction. The item did not give a third closed-response that some instructors preferred – which was to wait for the student or other classmates to notice the error themselves before correcting it. I also noticed a pattern in open responses whereby people who noticed the error that the limit notation had been dropped, tended to think the instructor should correct it, while instructors that did not notice what the error tended to lean towards moving on. Potentially due to these confounding issues, the item had a low item-rest correlation with the rest of the disciplinary obligation items (.19). The remaining item-rest correlations were .33 or greater. A possible improvement to the item would be to make the error more obvious, so that everyone would notice it and responses would reflect decisions about whether or not to correct it after it is noticed, rather than reflecting merely if it was noticed. It could also include a slide where the instructor asks if anyone wants to comment and waits, and no one responds – to eliminate the desire to wait for the students to point out the error instead of choosing one of the possible responses.

I dropped two additional items, A1155 and A104L, because they were redundant with each other and A102L, as indicated by correlations and modification indices. All three items had to do with going on an interesting mathematical tangent instead of staying with the planned mathematical content. Again, while it is useful in the development of a scale to include redundant items, the final version should take out ones that are excessively redundant (DeVellis, 2016). I recommend including these items in future iterations of the instrument, but with the caveat that they may eventually need to be removed in the analysis.

Table 5.9: Comparison of factor loadings for the total sample versus partial sample A for the PROSE-disciplinary instrument

Item	Loading for total sample, *n*=211	Loading for partial sample A, *n*=164	Difference
A1115	0.42	0.46	0.04
A1045	0.55	0.54	-0.01
A1125	0.66	0.65	-0.01
A1025	0.69	0.67	-0.02
A1145	0.69	0.67	-0.02
A1085	0.57	0.62	0.05
A102L	0.32	0.28	-0.04
A1015	0.56	0.62	0.06
A1065	0.56	0.56	0
A105L	0.51	0.51	0
A1105	0.55	0.54	-0.01
A101L	0.57	0.57	0
A106L	0.48	0.49	0.01
A103L	0.38	0.32	-0.06
A1095	0.64	0.68	0.04
A1055	0.54	0.49	-0.05
A1135	0.58	0.60	0.02

The final model did not have perfect fit as a unidimensional scale (RMSEA=0.07, CFI=0.93, TLI=0.92), with significant factor loadings ranging .32~.69. Examples of scenarios, item text, and item loadings or details as to why they were dropped are recorded in Appendix E; item loading comparisons between the complete and partial sample A are shown in Table 5.9.

PROSE-Institutional

The scenarios in the PROSE-institutional obligations instrument featured an instructor breaching some norm on account of adhering to department, university, or state policies. The recognition of the institutional obligation is measured by whether the instructor agrees that a breach of the contract on account of an institutional policy was appropriate or not. The items with highest loadings had to do with staying on pace with the other sections, staying focused on the material in the syllabus and that will appear on coordinated exams, and following department rules. Overall, the items in the PROSE-Institution instrument had uniqueness issues. Many items were too unique in their content to group with many or any other items in the set. In an EFA, all but one item had unique variances above .7, indicating that the items had high variance that was not explained by our targeted latent factors. This is reflected in the relative low loadings and high variance in the final model. A potential method to create more items that are not as disparate would be to create more items that targeted similar concepts as the items with highest loadings (e.g., instructing following the institutional guidelines associated coordinated courses like pacing and uniform exams). Or, if the content of a unique item involves an aspect of the institutional obligation that the researcher thinks is worth retaining, a set of items concerning variations of the same content could be developed.

There were three items, A2045, A2105, and A2155, that may have had issues due to participants' college-level contexts. The item-level statistics showed that A2045 did not hold well with the remaining items. The item A2045 was designed to probe if instructors would attend to the institutional obligation by giving up instructional time to allow some students to talk about student government issues. Many open responses reflected a desire to attend to the institutional obligation of supporting student government, but reasoned they still could do so through other means, such as sending the information by email, having the students speak before or after class, or asking the students to email to reschedule them at a time when the instructor can plan time for it. But these other means were not reflected in their closed-responses selections. The item consequently had a negative item-rest correlation (-.21). From what I gathered from open responses, student government did not carry the same importance at the university level as it might for the high school level. I think the item could be improved by making the issue be something more relevant to the college-level, for example students coming to talk about a campus visit day for future students or a undergraduate or graduate research conference run by the university. Then, the results would show whether the instructors would attend to the institutional obligation, given an issue that they know would be important to their college or university.

Item A2105 was intended to probe whether an instructor would honor a student's desire to remain in class or honor the institutional policy to have them attend a different mathematics class. In the open responses, many responses reflected that they would not even make this choice. In the college context, they said it would be up to the students' advisor to tell the student what class to take. Possibly due to this rejection of the scenario, the item had a poor loading (.24). A way to improve the item would be to have the decision be whether to say nothing, or to honor the institutional system in place by telling the student to check with their advisor if the course would be appropriate.

Item A2155 was about attending to the institutional obligation by complying with the policy of having a bilingual student take the test in another location with a translator, as opposed to staying in the same room with everyone. Most open responses reflected that, in a college context, this decision should be up to the students. They say that the instructor should make sure the student knows the resource is available, but not force the student to use them. As with other items where the participants say it should be left up to the student, the closed responses seem to be arbitrarily chosen in comparison to their corresponding open responses, perhaps explaining the low loading (.16). A way to improve this item would be to change the wording to "*encourage* the student leave in order to meet the requirements of the departmental policy" rather than the current wording of "insist the student leave in order…". Then the instructor would be able to leave the decision up to the student while still being able to attend to the institutional obligation by encouraging the department policy. There was another item (A2035) that that tlaso had issues of correspondence between closed and open responses.

Item A2035 was intended to capture if the instructor would recognize the institutional obligation by encouraging students go to a talk hosted by the university rather than remain in class. The vast majority of open response options stated that students are adults and should be able to make that decision for themselves (a theme that came up in the open responses of the last two items, as well). However, there was no consistent closed response option corresponding to this open response. There are examples of choices from 1 – strongly disagree to 6 – strongly agree chosen because the instructor thinks it should not be up to the instructor to decide. Thus, the item did not measure what was intended and had a negative item-rest correlation (-.21). I think this item does measure some aspect of the institutional obligation, namely whether instructors take it upon themselves to encourage student behavior that benefits the institution, but it is so far removed from the other instances where the institutional obligation applies that the item was too unique. Thus, the scale could be improved by generating more items on the topic of encouraging students to attend or participate in events sponsored by the university or department.

There were three items, A2075, A2135, and A206L5 that could benefit from making the role of the institution clearer in the content of the scenarios. Item A2075 featured an instructor who insisted that the students use the software provided by the mathematics department instead of allowing them to do the assignment without a computer. Open responses repeatedly reflected that participants agreed the students should use the software provided, not because it was required by the institution, but because it is valuable to learn to use new software. Although these responses corresponded high towards agreeing with the institutional obligation as instructors thought the instructor did the appropriate thing, the participants' open responses actually showed the reason for their choice had nothing to do with complying with university policy. This may explain why the item had a low loading (.24). The item could be improved by having the instructor say something like, “I think X software would be better for this activity, but we’re using Y software because the mathematics department says they want students to leave the program with the ability to use it.” Then, participants would have to choose between using a more appropriate software or honoring the institutional obligation to the mathematics department.

Item A2135 was meant to check if instructors would recognize the institutional obligation by returning to the regular curriculum given the presence of the department chair, as opposed to continuing with an extraneous example. Responses reflected that teaching should not change regardless of the presence of any outside evaluator, but they differed on how they felt about sticking with the regular curriculum versus the extraneous examples. The presence of the department chair seemed more like a red herring, where the true issue relevant to the institutional obligation was whether or not participants felt like sticking to the regular curriculum was important. An improvement to the item would be to remove the presence of the department chair, which hopefully would improve the item’s low loading (.13) by bringing participants’ attention to the actual issue relevant to the institutional obligation.

Item A206L5 concerned whether to help a few students who were falling behind in the course by providing them with a different assignment, as opposed to having the student complete the original problem set. In this scenario, having students complete the original problem set is supposed to indicate recognition of the institutional obligation, but it is not clear in the item that the original problem set is in any way associated with the mathematics department policy or required curriculum. The open responses reflect that participants were worried about issues like fairness for all the students, and making sure the students would be able to learn the same mathematical content from the different itemset. As this was a linking item, I did not have control over the content of this item in an effort to make it match the linking item in the other PROSE itemsets for K-12 level instructors. I did not use this item as the choice of having the student complete the original itemset was not clearly linked to recognition of the institutional obligation. As such, the item had a low item-rest correlation (0.00). The item could be improved by making the institutional connection more explicit. For example, the statement could be improved with the following wording, "The teacher should help the students complete the original set of problems with the required content, rather than provide them with a different assignment."

There were some items, A2065, A2145, and A203L5, that functioned as intended in terms of testing recognition of the institutional obligation, but were too unique to work well in a scale. Item A2065 was intended to test whether participants would attend to the institutional obligation of going over course information instead of using the time to go over a worksheet. The idea was that the mathematics department had requested that instructors review what mathematics courses students would need to take next for their various majors. The open responses reflected that the item functioned as intended. That is, some open responses said that it was not important to go over course information, as the students could get the same information from a handout or by talking with their advisor. Others said that it's important to respect department wishes. Some participants even criticized the department for making this a policy, but said they would still comply. I don't think the item needs to be revised, but there need to be more items created concerning department policies to use class time to cover topics that are not directly related to the mathematical knowledge at stake so that this item is not as unique. As is, it had a low loading (.17).

Item A2145 was about whether the teacher should recognize the institutional obligation by sticking with the class material, or use the time to talk about a historic topic. Most participants liked the idea of including the small history lesson, as long as it was short. I checked the summary statistics, thinking that this item would have overwhelming rejection of the institutional obligation, in favor of making time for the short story. However, the mean appropriateness rating was not extremely high (M=4.69). I am not sure why this item did not elicit the closed responses that corresponded to what they said in their open responses, but this disconnect could have contributed to the low loading (.09). I am not sure how to improve this item. However, a possible way to improve the scale would be to include more items like this so it would not be so unique. There were not many items in the PROSE-institution itemset that juxtaposed the institutional obligation with some sort of mathematics enrichment norm. Most items juxtaposed recognition of the institutional obligation against some norm to attend an individual student's idiosyncrasies for understanding the content.

Item A203L5 was intended to have the participant choose between recognizing the institutional obligation by asking a student to stop using a graphing calculator to comply with departmental policy or letting the student keep presenting at the board without interruptions. Most people thought that the instructor should not interrupt the student for fear of embarrassing the student or interrupting the train of thought. However, they thought the instructor should have checked beforehand or announced publicly that graphing calculators could not be used. There were some outliers who thought the policy should be followed no matter what. Overall, the item seemed to function as intended, but it had a low loading (.23). Thus, I would not completely discard this item, but not use it as one of the stronger indicators for the institutional obligation if I needed to worry about model fit or too many parameters.

Conversely, there were items that had the opposite problem. Items A2085, A2095, and A2185 had too much covariance. Similar to the previous three items, these items addressed the institutional obligation as intended, but their content did not contribute to the fit of the scale. They could be kept if the model fit and number of free parameters was not a high concern. Item A2085 had a high covariance with A201L. Item A2085 concerns whether the instructor should comply with the department policy to complete homework online or let the student turn in a hard copy of homework. Item A201L concerns whether the teacher should go over the topics in a set of review materials or ask students to review the topics on their own. I removed item A2085 because A201L is a linking item, meaning it is an item that is exactly the same for all levels of the PROSE instrument, and allows for comparison across levels. In an effort to respect the long-term goals for this scale, I kept A201L to maintain the ability to compare the college level instructor responses against the results from instructors of other levels (e.g., high school). However, it is more apparent to me how item A2085 elicits responses relevant to the institutional obligation than item A201L, so I could imagine an alternative (though not necessarily better) argument for dropping A201L instead.

Item A2095 had high covariance with A2075, A2185, and A202L. Item A2095 was about the appropriate response to a student who wanted to view a past final, given that the university policy was not to show past finals. Item A2075 was about doing an assignment without a computer, item A2185 was about delaying homework in a coordinated course to have a promised activity, and item A202L was about whether the instructor should answer all student questions, or table some of them to attend to the institutional obligation of keeping pace with the other sections of the course. The open responses to item A2095 showed that it generally elicited the intended response. That is, participants either said that the instructors should comply with university policy, thereby attending to the institutional obligation, or gave a reason for why they thought they could dismiss university policy in this circumstance. I had already decided to drop A2075, and A2185 had a high covariance with two additional items (A2055 and A205L5). I ended up retaining A202L because it was clearly about recognizing an institutional obligation and it was a linking item. I dropped item A2185, but similar to item A2085, the item tested the institutional obligation and there is reason not to automatically drop it if possible.

Finally, there are two items that did not function well because instructors either avoided the presented dilemma (A2015) or they said their choice depends on external factors (A2055). Item A2015 was designed to test whether instructors would honor the institutional obligation to clear the classroom after the class was over, or continue helping students who had questions. Many instructors agreed that the instructor should clear the class, but avoided the dilemma by saying that they could work with students in office hours or at the start of the next class. Instructors also mentioned being concerned for the students in the next class and their ability to start on time and have a fair experience, which sounds more like recognition of the obligation to the individual than to the institutional policy of ending class on time. This item did not cohere with the rest of the items, as it had a negative item-rest correlation (.-17). The item might function more as intended by adding a slide where the teacher says, “I don’t have any more office hours this week, but I can stay a few minutes to finish answering questions.” Then participants would be forced to choose between recognizing the institutional obligation by clearing the class or answering the students’ questions, instead of feeling like they could attend to both at once.

Item A2055 was designed to measure instructors' recognition of the institutional obligation by asking whether an instructor would allow a student to stay in the class even though they missed the registration deadline. The open responses from instructors often indicated that their closed-response decision depended on whether or not they thought they would have any power to override the enrollment deadline. The item became less about whether or not they would like to honor the enrollment deadline, but whether or not they thought it would be possible to influence it. Other responses interpreted the item as asking whether or not they would allow the student to sit in the class unenrolled or as an audited class, so they agreed with the item in their closed response, but without recognizing the institutional obligation as designed by officially enrolling the student in the course. This item had a low loading (.172) and could be improved by clarifying that the instructor did have the power to override the deadline.

Table 5.10: Comparison of factor loadings for the total sample versus Partial Sample A for the PROSE-institution instrument

Item	Loading for total sample, *n*=217	Loading for Partial Sample A, *n*=168	Difference
A202L	0.54	0.53	-0.01
A201L	0.36	0.37	0.01
A205L5	0.50	0.45	-0.05
A2175	0.33	0.26	-0.07
A2115	0.60	0.54	-0.06
A2165	0.42	0.52	0.10
A204L	0.41	0.32	-0.09
A2025	0.48	0.50	0.02
A2125	0.43	0.53	0.10

The unidimensional model without the items with negative item-rest correlations yielded poor fit (RMSEA=.08, CFI=.65, TLI=.60). I removed items with loadings <.1, and the model was only slightly improved (RMSEA=.08, CFI=.76, TLI=.72). After removing all the items outlined above, I arrived at a model with acceptable fit (RMSEA=.07, CFI=.92, TLI=.90) with significant loadings ranging .33~.54. The nine items in the final model, except perhaps A2125, deal with issues that could be present in a coordinated course (e.g., pacing, reviewing for an exam, covering what is on the syllabus). Loadings for partial sample A are only trivially different, as shown in Table 5.10.

Reliability

The internal consistency of the scale for each construct is reported for the item groupings developed in this chapter. Though I ran reliability statistics before conducting the factor analyses, I report the scores after so that I can base them on the final items used for each latent construct. Commonly used measures of internal consistency are mean inter-item correlations (IICS), Cronbach's alpha, and the factor rho coefficient (Kline, 2016). A common guideline for IICs is to consider values between .15 and .50 as acceptable (Clark & Watson, 1995), for Cronbach's alpha is to consider values over .7 as acceptable and those below .5 as unacceptable (Kline, 2016), and for rhos to be above .7 (Hair, Black, Babin, Anderson, & Tatham, 2006).

Table 5.11: Reliability statistics for the lower and upper-division versions of the INQUIRE instrument

Triangle Relationship	Constructs	IIC	α	Rho
Student-Content	Construct/Critique	.48	.90	.90
	Prove	.45	.87	.88
	Solve open problems	.38	.71	.71
Teacher-student	Interactive Lecture	.25	.62	.59
	Hint without tell	.38	.65	.69
Student-Student	Group work	.58	.90	.92
	Presentations	.54	.87	.87

Table 5.12: Reliability statistics for the latent constructs in the beliefs and PROSE instruments

Instrument	Constructs	IIC	α	Rho
Beliefs	Struggle	.38	.75	.75
	Model	.31	.76	.76
	Awareness	.42	.74	.80
PROSE	Individual	.19	.74	.74
	Interpersonal	.26	.86	.87
	Disciplinary	.25	.85	.85
	Institutional	.18	.66	.66

The fit for scales in the three instruments are satisfactory. Most item groupings had acceptable IIC. alpha and rho scores, as shown in Table 5.11 and Table 5.12. There are a few instances where the fit statistics are not acceptable (interactive lecture and hinting without telling alpha and rho scores, group work and presentation IICs, institutional obligation alpha and rho scores). It would be possible to make the fit statistics perfect by removing more items. However, I chose not to do that because the items themselves loaded well and could be justified by appeal to theory. As much as possible, I prioritized theoretical validity over meeting arbitrary cutoffs perfectly. The good fit of these scales has implications for how they can be used.

Consequences of Scale Development

The fit of the models and reliability statistics show that the scales are ready for use and that the groupings of items into constructs was conceptualized well. The INQUIRE instrument showed item loadings in the factor analyses that were almost equal. This shows that the items were conceptualized well, and that the scales could be used with classical test theory methods (assuming all items have equal loadings) with almost identical results to results using structural equation modeling. This is useful for the following chapter where I use a multinomial logistic regression or for future studies that may have smaller samples. Reliability indicators, rho and alpha, are calculated in similar ways, except that the rho statistic penalizes scales with redundant items. The similarity of the alpha and rho reliability statistics for each construct show that the care in the scale development to remove redundant items was effective.

The development of the PROSE and beliefs instruments with this sample enables the field to leverage the work on mathematics teacher education to college-level research. The results of the scale development have implications for the future use of each of the instruments. The PROSE instruments can now be used with the college population in addition to the K-12 populations from the original design. There are items within each of the four PROSE instruments which this study has shown could be improved, if the user would like to use the full set of items. The constructs theorized by Clark and colleagues (2008) are also applicable to the college level. There are more improvements to the INQUIRE instrument aside from revising the particular items that did not function well that I would like to incorporate in future iterations.

Future INQUIRE Iterations

The INQUIRE instrument was designed to be organized around the relationships depicted in the instructional triangle. The grain size of latent constructs within each relationship differed. While the constructs within the student-content and student-teacher categories pinpointed specific activities, the constructs within the student-student category (group work and student presentations) were broader activity structures, and the construct of teacher preparation could include a wide variety of broad or specific practices. During the group work and student presentation activity formats, there are not only student-student interactions – students are also interacting with the content, the teacher is interacting with the content that the students are producing, and the teacher is walking around interacting with students. The current items in the group work and student presentation scales do not attend to higher-inference nuances in instruction for the student-student interactions. For example, I noted earlier that including group work may be unrelated to the interpersonal obligation because fostering a productive social environment is more concerned with *how* the instructor manages group work or a whole-class discussion, rather than the class format itself. On the student-student level, I would add constructs to address how the instructor scaffolds the interaction, rather than only asking about the general formats of interaction.

A consequence of the teacher preparation construct being so broad and varied was that the items did not group together well as a unidimensional factor, as there were many other subconstructs, each with not enough items. I say in the scale development section that future versions of the instrument could include more items on each of the three constructs that the EFA revealed: examining student thinking, using student work, and creating opportunities for discovery. These are all key aspects of *instructional exchanges*, the "interpretive acts that the teacher needs to do" (p. 606) to make the mathematics at stake visible (Herbst & Chazan, 2012), such as introducing a new item of knowledge or giving a mathematical label to a sequence of steps. In the didactical contract, the instructor is responsible for managing these exchanges, but I conjecture that what the exchanges consist of changes given the context of inquiry-oriented practices. Some instructional exchanges that describe ways the instructor can interact with the content in a way that supports inquiry, based on preliminary explorations from the EFA, include the work of examining student thinking and student work to be able to respond in a way that guides students to the knowledge at stake, and crafting the tasks for students to do that will reveal the knowledge at stake.

Table 5.13: New constructs for future iterations of the INQUIRE instrument

Triangle Relationship	**Construct**	**Definition**	**Examples**
Student-Student	Students as teachers	Asking students to explain something to other students (could be during group work, student presentations, or during a lecture)	How often do you ask students to explain something that another student is struggling to understand? How often do you refrain from explaining something and ask a student to do so instead?
	Joint problem-solving	Prompting students to solve problems or construct proofs together (could be during group work, student presentations, or during a lecture)	How often do students consult with each other while working during class? How often do you have students share ideas with each other about how to solve or prove something?
Teacher-Content	Examining student thinking	Examining the mathematical content that students have produced during class to understand how they are thinking	How often do you read line-by-line through students' work during class to understand their reasoning? How often do you try to understand how people are arriving at incorrect solutions by looking at their work?
	Using student work	Looking at student-produced content to find useful examples	How often do you look for student work to illustrate a common misconception? How often do you examine what your students are doing to find an example to show to other students?
	Creating opportunities for discovery	Giving students scaffolds that will guide them to discover something	How often do you give student real-life situations to guide their discovery of some mathematical technique? How often do you design simple cases so students can discover more complex or general cases of the same phenomena?

The construct of teacher preparation (e.g., planning lessons and writing exams) differs from the other constructs in the instrument in that it's the only construct that accounts for activities that occur outside the classroom, and the INQUIRE instrument currently focuses on classroom practices. For the sake of consistency, the instrument might also need to account for other outside classroom practices, such as student interactions with the content during homework or teacher-student interactions over email and during office hours. In the next iteration of the INQUIRE instrument, I would replace the constructs of group work, student presentations and teacher preparations with the constructs in Table 5.13.

The most neglected part of the instructional triangle, by the INQUIRE instrument, is the environment. Some aspects of the environment were considered through use of the PROSE instrument, which accounts for the instructors' professional obligations. But a different aspect of the environment, that is an essential part of the model for inquiry, is creating an equitable environment for students to learn. Thus, future iterations of this instrument would have constructs designed to target issues of equity. For example, one construct in this category could be questioning whether the instructor practices equitable habits of mind . Yong (2018) encouraged college instructors to habitually ask themselves three questions when making instructional decisions: "(1) Who is likely to benefit? Who might not? (2) Who might feel included or excluded? (3) How would I know if I need to intervene?" Or, if these are not considered to technically be components of IOI, a separate instrument could be made to be used in conjunction with the INQUIRE instrument to examine interactions between the different practices and equitable habits. This separate instrument focused on equitable habits might fit as a new subset of the PROSE-interpersonal instrument. To offer a conjecture, perhaps more equitable habits of mind are used when deciding to have students present than when deciding to whom to “hinting without telling.” The complex issue with building such an instrument is that “equity” is not a one-dimensional construct, and mathematicians or mathematics educators can disagree on which concrete practices contribute or do not contribute to an equitable environment. But in future versions of the INQUIRE instrument, those issues would be worth investigating.

Chapter 6 **Results**

I respond to the research questions by first reporting what IOI practices instructors professed to use and what patterns of practices emerged in their use as discovered through a hierarchical cluster analysis. I answer the second research question about what factors predict IOI in two ways. First, because there were three patterns of inquiry-oriented practices, I conducted a multinomial logistic regression to understand what factors predict different patterns of IOI. Second, I created a structural equation model for each targeted practice of the INQUIRE instrument to further understand the relationship between each factor and individual practice. I report background characteristics of the instructors belonging to each cluster to explore what other factors might contribute to the use of inquiry-oriented practices.

What IOI Practices Are Used

To respond to the first research question of what practices are used, instructors reported using all the practices identified in the literature review and included in the INQUIRE instrument, but to widely varying degrees. The descriptive statistics from the lower-division INQUIRE instrument are shown in Table 6.1, where the values were self-reported following the 6-point scale: 1 - Never, 2 - A few classes, 3 - Less than half the classes, 4 - More than half the classes, 5 - Every or almost every class, or 6 - Multiple times per class. The most frequently used practices were making lecture more interactive and giving hints rather than telling, while the least frequently used was including student presentations. In the context of lower-division mathematics, I expected the low rates of student presentations due to the larger class sizes and uniform curricula, and that is what I found (M=1.81, SD=.93). The high means for interactive lecture (*M*=4.38, *SD*=.86) and hinting without telling (*M*=4.29, *SD*=.98) indicate that these might be the most feasible practices for instructors to implement at the lower-division level for classes like calculus. Considering that the sample includes all mathematics instructors, whether or not they profess to use IOI, I was not expecting the mean for group work to be as high as it is (*M*=2.98, *SD*=1.22). I expected many of these practices to be used more in upper-division courses, but rarely in lower-division courses.

Table 6.1: Mean and standard deviations for the lower-division INQUIRE instrument

Relationship	**Constructs**	***n***	**Mean**	**Standard Deviation**
Student-Content	Constructing/critiquing	252	2.03	0.86
	Proving	252	2.58	0.97
	Open problems	252	3.22	0.95
Teacher-Student	Interactive lecture	257	4.38	0.86
	Hinting without telling	257	4.29	0.98
Student-Student	Group work	194	2.98	1.22
	Student Presentations	194	1.81	0.93

Descriptive statistics for the upper-division INQUIRE instrument are shown in Table 6.2. The table also includes the lower-division descriptive statistics for the same set of instructors so that differences can only be attributed to the level of course, rather than characteristics of a different sample. For example, if the comparisons were not restricted to the same sample, differences in practice use could be attributed to the higher average teaching experience of those instructors that teach upper-division courses. I conducted a paired sample two-tailed *t*-test for each construct to check if differences were significant. As expected, instructors reported using almost all inquiry-oriented practices significantly more often in upper-division courses than lower-division courses, as shown in the right-most column of Table 6.2. The largest difference is seen in the practice of providing students with opportunities to prove and construct formal arguments ($t(77) = -9.07$, $p<.0001$), which is expected given that proofs become the central activity beginning in upper-division courses.

The only practice that did not occur significantly more in upper-division courses was giving students hints without telling them how to proceed ($t(77)=-1.59$, $p=.12$). The next smallest difference between upper- and lower-division courses was for the practice of using interactive lecture ($t(77)=2.08$, $p=.04$). A potential explanation for this pattern is that these two practices are arguably the most feasible to implement without making large changes to the content and format of the course. The forthcoming structural equation models give further insight into why instructors profess to use these two practices with greater levels of frequency in lower-division courses. In addition to seeing what IOI practices instructors say they used, I also explore how they say they use practices in relation to each other. Instructors do not only vary widely on which practices they report using, they also vary by the patterns of practices they report using for the same course. I investigate this with a cluster analysis.

Table 6.2: Mean, standard errors (in parentheses), and comparison t-*test results for the INQUIRE instrument,* n=*78*

Relationship	**Constructs**	**Lower-Division Courses**	**Upper-division Courses**	**Difference**
Student-Content	Constructing/critiquing	2.12(.11)	2.75(.13)	-.63(.08)***
	Proving	2.65(.13)	3.81(.14)	-1.15(.13)***
	Open problems	3.21(.12)	3.80(.12)	-.60(.13)***
Teacher-Student	Interactive Lecture	4.49(.09)	4.30(0.11)	.19(.09)*
	Hinting without telling	4.15(0.12)	4.33(.14)	-.19(.12)
Student-Student	Group work	2.94(.15)	3.19(.16)	-.24(.12)*
	Student Presentations	1.68(.10)	2.10(.13)	-.42(.10)**

*$p<.05$, **$p<0.01$, ***$p<0.001$

Cluster Analysis

I explored what groups of instructors used similar patterns of inquiry-oriented practices using a hierarchical cluster analysis. Using Ward's method to minimize within-group variation, I generated and viewed dendrograms for 1-12 groupings. The examples in Figure 6.1 and Figure 6.2 show the groupings for N=12, where N represents the number of groups the algorithm first generates. The initial N chosen is unimportant, so long as it shows enough information to be able to see a pattern. The dendrograms indicate that there are two all-encompassing clusters of instructors – those that tend to use inquiry-oriented practices (on the left) and those that do not (on the right). The dendrograms further illustrate that, within those groups, there are nonuniform subgroups. I added a horizontal line to Figure 6.2 to show where the groups could be cut off due to the natural break in the heights of the groups. This would break the groups into two – one group that does more IOI practices and one group that does less. However, I chose the cutoff location shown in Figure 6.1, at the next level down than in Figure 6.2, because I was interested in the characteristics of the subgroups of people that used more inquiry-oriented practices and the characteristics of the subgroups of the people did that used less inquiry-orietned practices. Thus the dendrograms revealed four clusters of inquiry-oriented practice use – with two groups that use different patterns of inquiry-oriented pratices with relative high frequency, and two groups that use them with relative low frequency.

Figure 6.1: Dendrogram for N=12 groups using Ward's method to minimize within-group variation of inquiry-oriented practice use, with cutoff line

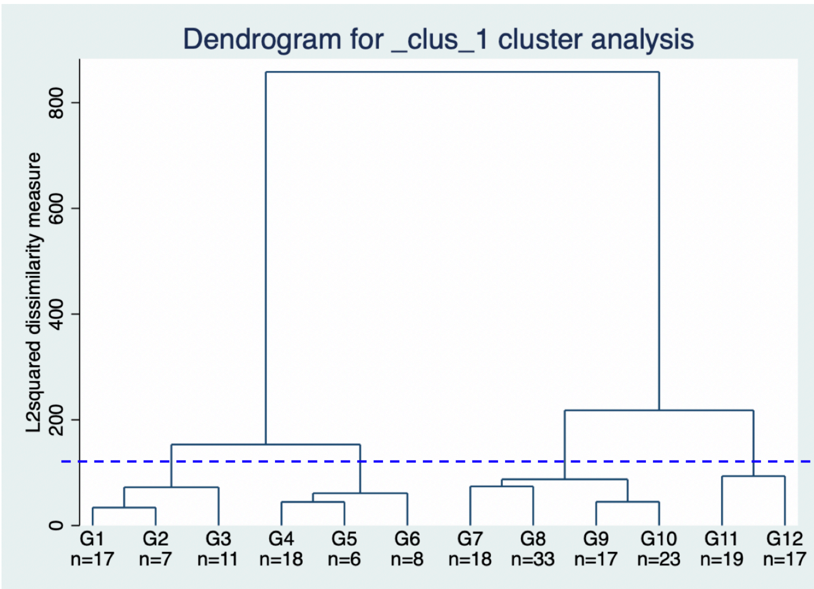
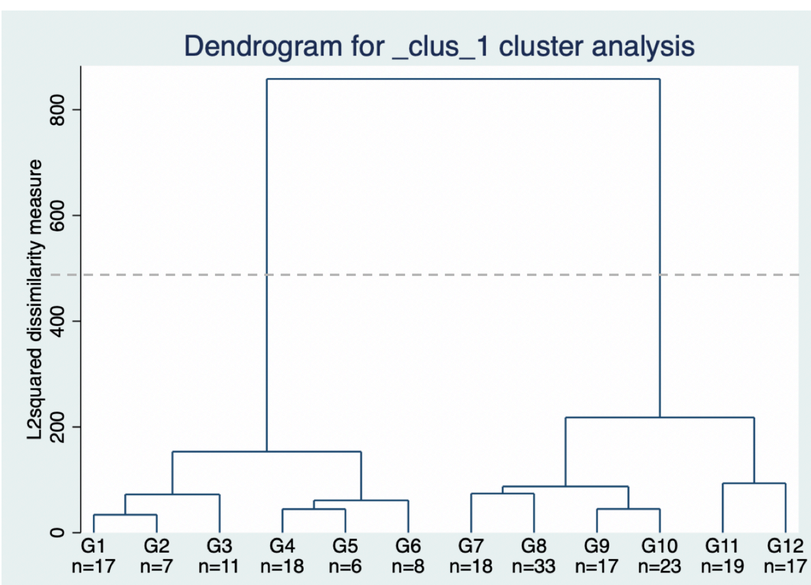


Figure 6.2: Dendrogram for N=12 groups using Ward's method to minimize within-group variation of inquiry-oriented practice uses, with a potential (not used) cutoff line



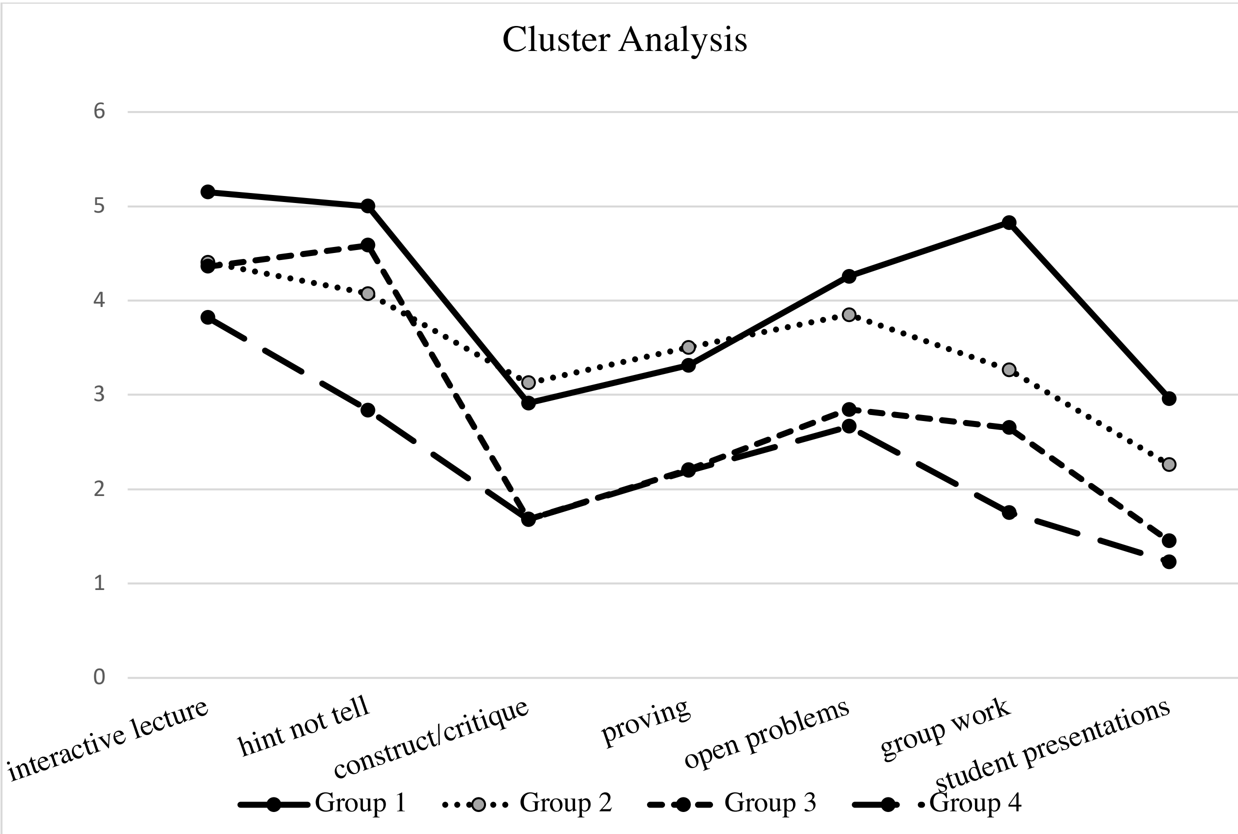
The mean values of inquiry-oriented practice use for the four clusters are given in Table 6.3 and illustrated in Figure 6.3. The instructors in the first cluster reported using the practices of interactive lecture, hinting without telling, giving students opportunities to solve open problems, having students work in groups and having students give presentations to the class the most. It could be characterized as using most of the inquiry-oriented practices relatively often. They provide students with frequent opportunities to work in groups and work on open problems, in comparison to the other groups. The second cluster of instructors could be characterized by providing opportunities for students to engage deeply with mathematical content, through activities that involve proving and constructing and critiquing claims. This second cluster bears some semblance to the Moore Method, with higher attention to giving students challenging content, and less emphasis than the first cluster on group work. The instructors in the third cluster reported emphasizing the practices that involve interactions between the instructor and student through interactive lecture or hinting and not telling when responding to students, but used the rest of the practices at a relatively low frequency. The third cluster was, by far, the largest cluster (*n*=91), indicating that the two practices of making lecture more interactive and hinting without telling might be more accessible to instructors who might otherwise be uninterested in using IOI. The instructors that composed the fourth cluster reported using all inquiry-oriented practices to a lesser extent.

Table 6.3: Mean values for each inquiry-oriented practice divided by clusters

Cluster	*n*	active lecture	hint not tell	construct/ critique	prove	open	group work	student presentations
1	35	**5.15**	**5.00**	**2.91**	**3.31**	**4.26**	**4.83**	**2.96**
2	32	**4.41**	4.07	**3.13**	**3.50**	**3.85**	**3.27**	**2.26**
3	91	4.36	**4.59**	1.68	2.21	2.85	2.65	1.45
4	36	3.82	2.84	1.68	2.20	2.67	1.75	1.23

The two highest means for each construct are in bold.

Figure 6.3: Bar graph of means for each inquiry-oriented practice, separated by cluster



This cluster analysis gives evidence that there is not one version of IOI. Although there is a clear separation between instructors that do some version of IOI and those that do not, instructors' reports illustrate that there is variation among patterns of implementation. These three clusters of IOI practitioners give researchers insight to the versions of IOI that exist.

Predicting Use of Inquiry-oriented Practices by Cluster

The multinomial logistic regression analysis showed that beliefs and professional obligations significantly predicted membership in the clusters of inquiry-oriented practices, as shown in Table 6.4. The regression compared membership in each of the three clusters of patterns of IOI practice use against the fourth cluster, which was the group characterized by the least use of every inquiry-oriented practice. Recall that the first cluster had high frequencies in all inquiry-oriented practices relative to other clusters, especially by providing students opportunities to work in groups. I found that instructors' belief that students should struggle, instructors' awareness of students' dispositions, more recognition of the individual obligation, and more recognition of the institutional obligation all significantly predicted membership in this first cluster, with respect to the fourth cluster.

Table 6.4: Multinomial logistic regression predicting cluster membership with respect to the fourth cluster, n=*159*

		Cluster 1		**Cluster 2**		**Cluster 3**	
Instrument	Predictors	Coeff	SE	Coeff	SE	Coeff	SE
Beliefs	struggle	2.35***	0.61	1.37**	0.51	0.93*	0.41
	model	1.15	0.63	0.28	0.55	0.78	0.44
	awareness	1.78***	0.44	1.12**	0.39	0.67*	0.29
Professional Obligations	individual	2.36***	0.72	0.17	0.61	0.16	0.46
	interpersonal	-0.14	0.88	-0.42	0.78	0.06	0.58
	discipline	-0.3	0.82	-0.6	0.74	-0.07	0.56
	institution	1.33*	0.68	-0.75	0.57	0.35	0.43
	_cons	-30.01	6.81	-3.98	4.62	-8.53	3.87

*p≤0.05, **p≤.01, ***p≤.001

The recognition of the individual obligation significantly predicted membership in the first cluster. If a participant had one more unit of recognition of individual obligation, their relative risk ratio (i.e., odds of belonging to cluster 1 as opposed to cluster 4) would be $e^{2.36} = 10.59$. I hypothesize that this high ratio is because an instructors' willingness to deviate from the expected didactical contract for college undergraduate mathematics to meet their individual students' needs is closely related to their willingness to deviate from the contract to meet the students' needs in terms of learning the material. Some of the highest indicators of recognition of the individual obligation from the PROSE instrument had to do with adjusting lessons to engage a particular student, grading a student's quiz based on what the student was able to finish rather than the normative way of assigning zero points to problems left blank, and letting a student use a special learning aid. The type of instructor willing to breach norms to engage particular students would likely also be willing to, for example, engage students in practices like group work or making lecture more interactive.

The institutional obligation also predicted membership in the first cluster, but for less obvious reasons. If a participant had one more unit of recognition of the institutional obligation, their relative risk ratio (i.e., odds of belonging to Cluster 1 as opposed to Cluster 4) would be $e^{1.33} = 3.79$, if all other variables were held constant. That is, holding a higher recognition of the institutional obligations would make an individual more likely to belong to cluster 1 than cluster 4. The items that were strong indicators of recognition of the institutional obligation in the PROSE instrument had to do with keeping with the pace of other sections of a coordinated course, and teaching to prepare students for upcoming exams. Thus, I had hypothesized that the institutional obligation would be inversely related to the membership in Cluster 1 as compared to Cluster 4. A possible explanation for the direction of this relationship is that mathematics departments are advocating for more active learning due to all the institutional pressures from stakeholders like organizations and parents. So when instructors recognize the institutional obligation, it directly implies using inquiry-oriented practices because their departments have explicitly recommended it. An alternative conjecture was that, because graduate student teaching assistants (GTAs) make up a large portion of the sample, they have higher recognition of the institutional obligation due to the nature of their position, and thus are more disposed towards trying inquiry-oriented practices. However, controlling for the variable of being a GTA did not alter the significance of the individual or institutional obligations. In the following sections I look at which practices are predicted by the institutional obligation and which are not, to give a better sense of what might explain this relationship.

Instructors in the second cluster emphasized the practices that provided opportunities for students to engage deeply with mathematical content, through activities that involve constructing and critiquing claims and proving. The belief that students should be given opportunities to struggle and awareness of students' mathematical dispositions significantly predicted membership in cluster 2, compared against cluster 4. None of the professional obligation predictors significantly predicted membership in this cluster. I hypothesized that the disciplinary obligation would have a significant positive relationship with this cluster, but it did not. The items in the PROSE instrument having to do with the disciplinary obligation did not focus heavily on having the students engage with the mathematical content, but rather had to do with representing mathematics appropriately and elegantly, showing students applications, or expanding on mathematical theory – only 3 of the 17 items in the final scale had to do with asking students to make generalizations. So perhaps a different perspective of what the disciplinary obligation entails would have been significantly related to this cluster. Another explanation for the lack of significance of all the professional obligations is that, while cluster 2 is different from cluster 4, it is not extremely different. This explanation is also applicable to the results for cluster 3.

Cluster 3 included instructors who did not tend to say they were doing practices much different from cluster 4, except they are higher on the practices having to do with student-teacher interactions (i.e., the practices of making lecture more interactive and hinting without telling). Instructors' belief that students should have opportunities to struggle and awareness of students' dispositions positively predicted membership in cluster 3. However, recognition of the professional obligations did not significantly predict membership in cluster 3. I turn to reporting characteristics of people within each group to see what else could explain membership to each cluster.

Cluster Characteristics

The background characteristics for each group from the cluster analysis are given in Table 6.5 and Table 6.6. I suspected that, in the groups with more IOI (groups 1 and 2), it would be more likely that a given participant would be female, have less faculty research obligations, teach smaller courses, and claim to use IBL, in comparison to participants in the group that professed to do IOI practices with the least frequency (group 4). Few of these patterns ended up being apparent. It does appear that graduate students tend to have a higher presence in groups 1 and 2. The higher percentage in group 2 makes sense, considering that group 2 has the highest use of the practices of giving students opportunities to construct, critique and prove, but not as high use of all other practices as Group 1. I could imagine graduate student teachers giving students difficult mathematical work to struggle with, perhaps with less of the other in-class scaffolding like time to work with peers or interact during the lecture.

Another pattern of note is that while there are more people that profess to use IBL in groups 1 and 2, as shown in Table 6.6, the distinction is not as marked as one might would suspect. This reinforces the original motivation for this study, that there is not a single way to define what IOI or IBL is. Moreover, studies that simply ask participants to self-report whether or not they do IOI or IBL likely have validity issues due to different interpretations of that word.

Table 6.5: Basic background characteristics in percentages of instructors in each IOI group from the cluster analysis

		Group 1	**Group 2**	**Group 3**	**Group 4**
Mean years of experience (SD)*		7.91 (7.49)	5.19 (5.39)	6.96 (6.71)	9.64 (8.79)
Gender	Female	42.86	64.52	44.94	57.14
	Male	54.29	35.48	51.69	42.86
	Another identity	0	0	1.12	0
	Prefer not to answer	2.86	0	2.25	0
Position	Graduate student	54.29	61.29	46.59	35.29
	Postdoctoral researcher	5.71	9.68	11.36	5.88
	Part-time faculty	2.86	0	6.82	2.94
	Non-tenure-track faculty	20	22.58	27.27	26.47
	Tenure-track faculty	11.43	3.23	5.68	14.71
	Tenured faculty	5.71	3.23	2.27	14.71
	None of the above	0	0	0	0

*Only category not given in percentages

Each of the background variables were added to the multinomial logistic regression in the previous section to test their significance. The variables that were not ordinal were dichotomized. Though there are arguably some visible patterns in the table, none of the background variables were statistically significant when included in the multinomial logistic regression in the previous section. The only background variable with a *p*-value below .3 was *graduate student*, with a coefficient of 1.85 (*p=0.096*) in predicting membership to the first group. It did not predict membership to the second group, as I thought I had observed from the descriptive statistics. This result is interesting because it suggested that inexperienced instructors may be the most open to using inquiry-oriented practices, despite often being restricted by more institutional constraints guiding their instructors.

Table 6.6: Background characteristics related to research and teaching in percentages of instructors in each IOI group from the cluster analysis

		Group 1 (SD)	Group 2 (SD)	Group 3 (SD)	Group 4 (SD)
Carnegie classification	Research I	65.71	64.52	70.79	68.57
	Research II	25.71	25.81	15.73	22.86
	Research III	2.86	9.68	1.12	0
	Unsure	5.71	0	12.36	8.57
Research	Yes, they do research	77.14	74.19	75.28	74.29
	No, they do not do research	22.86	25.81	24.72	25.71
Papers published	0	58.82	67.74	62.92	60
	1	14.71	19.35	17.98	20
	2	14.71	12.9	14.61	14.29
	3+	11.76	0	4.49	5.71
Research to teaching ratio	0-25% research, 75-100% teaching	28.57	41.94	46.07	42.86
	25-50% research	25.71	16.13	15.73	25.71
	50-75% research	37.14	35.48	21.35	17.14
	75-100% research	8.57	6.45	16.85	14.29
Number of courses teaching	1	40	32.26	40.45	37.14
	2	34.29	48.39	30.34	37.14
	3	20	6.45	21.35	20
	4+	5.71	12.9	7.87	5.71
Used IBL	Yes	46.67	50	35	31.82
	No	33.33	50	47.5	54.55
	Unsure	20	0	17.5	13.64
Enrollment for smallest course usually taught	<25	45.71	35.48	37.08	40
	26-45	51.43	48.39	50.56	45.71
	46-100	2.86	16.13	7.87	11.43
	100+	0	0	4.49	2.86

The main two weaknesses in the representativeness of this study's sample are that it has an overrepresentation of females and graduate student instructors compared to the national population. The lack of significance in females and graduate students predicting group membership indicates that there is no reason to believe that the greater proportion of females in the sample has altered the existence or membership of these ways of practicing IOI from the known characteristics of university mathematics instructors. Though not significant, the pattern of being a graduate student heightening the probability of belonging to the first cluster could indicate that there are fewer members of Group 1 in reality than indicated by this study. Given these patterns of practice, I turn to understanding if and how beliefs and professional obligations can predict the individual practices that these clusters of practices are composed of. Parsing the analysis could help give a better sense of how beliefs and professional obligations relate to each of these practices individually.

What Factors Predict Inquiry-Oriented Practices

I created separate structural equation models for each inquiry-oriented practice. I used structural equation models instead of OLS regression because not all the items were equal indicators of the latent construct they were predicting for all the instruments. I modeled each practice separately for theoretical and practical reasons. Theoretically, the study is aimed to investigate how beliefs and obligations predict specific practices rather than how practices are related to each other. Practically, an ideal minimum sample size would be ten times the number of parameters estimated (Bentler & Chou, 1987). A model with one practice and the beliefs and obligations contains 103 free parameters, and a model with all practices would contain 220 free parameters, which is nearly the number of participants in the total sample. Newer studies show that the minimum sample size depends on many different factors and could be smaller than previous cutoffs recommended (e.g., Wolf, Harrington, Clark & Miller, 2013), but still not small enough to enable all the practices to be modeled at once with the data collected in this study. For these analyses I use the results from the lower-division item set of the INQUIRE instrument to utilize a greater sample size (for the total sample, n=257 as opposed to n=78).

Table 6.7: Fit statistics for each of the structural equation models included in the results

Model that predicts	CFI	TLI	RMSEA
Construct/Critique	.94	.93	.05

Student-Content	Prove	.94	.93	.05
	Open problems	.91	.90	.05
Student-Teacher	Interactive lecture	.93	.91	.05
	Hint not tell	.93	.92	.05
Student-Student	Group work	.94	.92	.05
	Presentations	.93	.92	.05

The models targeting each the inquiry-oriented practices fit well, as shown in Table 6.7. In general, these fit statistics show that beliefs and professional obligations are acceptable indicators for modeling the inquiry-oriented practices that instructors choose to implement.

Student-Content Models

The results from structural equation models predicting the student-content inquiry-oriented practices showed some of the hypothesized relationships between beliefs, professional obligations, and practices (see Table 6.8). The practice of giving students more opportunities to construct and critique was predicted by the awareness of students' mathematical dispositions (ß=.388, $p<.001$) and the individual obligation (ß=.35, $p<.01$). If an instructor is concerned for the needs of individual students, it logically follows that they also could care about giving students ways to engage with the mathematics. Mahavier (1999) contended that a main reason people who try to teach using inquiry-oriented practices find themselves without success is lack of patience for individual students. He continued that it takes a lot of patience to let students figure things out for themselves. If an instructors' beliefs show a tendency towards using more of this practice, recognition of the individual could be additionally leveraged as an asset towards successfully implementing the inquiry-oriented practice of constructing and critiquing.

The practice of giving students more opportunities to prove was predicted by beliefs but not by recognition of any of the professional obligations. The belief that students should be allowed to struggle was positively correlated to giving students more opportunities to prove (ß=.31, *p*<.05), as would be expected. The practice of giving students opportunities to prove was also predicted by instructors' awareness of students' mathematical dispositions (ß=.24, *p*<.05). If an instructor invests the time to understand how the student relates to the mathematics in their everyday lives (as is asked in the items on awareness of mathematical dispositions), the instructor is also likely to be one that gives their students things to prove.

Table 6.8: Results from structural models predicting the student-content latent constructs, n=265

		Construct/Critique		**Prove**		**Open Problems**	
Instrument	Predictor	Standardized Estimate	SE	Standardized Estimate	SE	Standardized Estimate	SE
Beliefs	Struggle	.26†	.15	.31*	.16	.36*	.17
Professional Obligations	Model	-.01	.17	.07	.18	-.18	.19
	Awareness	.38***	.08	.24*	.09	.32***	.10
	Individual	.35**	.13	.08	.14	.42**	.16
	Interpersonal	-.20	.14	-.19	.16	-.21	.17
	Disciplinary	-.25	.15	-.12	.17	-.25	.18
	Institution	.10	.10	.05	.11	.37**	.12

† *p*<0.10, **p*≤0.05, ***p*≤.01, ****p*≤.001

The practice of presenting students with *open problems* to solve was significantly predicted by instructors' belief that students should be allowed to struggle (ß =.36, *p*<.05), awareness of students' mathematical dispositions (ß =.32, *p*<.01), the individual obligation (ß = .42, *p*<.01), and the institutional obligation (ß = .37, *p*<.01). The two beliefs and obligations have positive coefficients, indicating that if an instructor holds more of those beliefs and recognizes those obligations, they are more likely to give students open problems. Similar to the significance of the recognition of individual obligation for the practice of constructing and critiquing, it might take the same type of instructor who has the patience to attend to students' individual needs who is also willing to attend to students' individual, unique solutions and ways of working through open problems. I am unsure as to why the institutional obligation was significantly related to this practice as opposed to the other two student-content practices. One possible explanation is that, in lower-division courses, open problems might be part of the curriculum more often than problems with proving or constructing and critiquing claims. For instructors feeling institutional pressure to incorporate active learning, giving students open problems may seem more useful to convey the mathematical content than having students at that level construct or critique claims or write proofs. In many mathematics departments, there is a gateway or transitions course to introduce proofs, so instructors may feel hesitant to try to teach it beforehand.

I initially thought that another possible explanation for the positive association between the recognition of the institutional obligation and the use of open problems could be due to the high number of graduate student teaching assistants (GTAs) in the sample. Perhaps GTAs are more inclined to use open problems, and also tend to attend to institutional obligations more due to the lack of autonomy they hold in their teaching positions, relative to other instructors. However, I tried controlling for the characteristic of being a GTA, and the recognition of the institutional obligation was still significant (with a slightly smaller *p*-value). In fact, using a two-variable OLS regression, being a GTA negatively predicts recognition of the institutional obligation (ß =-.25, $p<.007$). More work could be done to understand why the institutional obligation performs this way.

In addition to noting the significant relationships between beliefs and obligations and inquiry-oriented practices, it is also worth noting that some constructs are not significant predictors of these practices. Specifically, the interpersonal and disciplinary obligations are not significant predictors of inquiry-oriented ways of engaging students with the content. I had hypothesized that the disciplinary obligation would predict a higher use of these practices because someone who wanted to engage students with authentic mathematical experiences would want to give students challenging opportunities to do mathematics beyond algorithmic procedures. However, I noticed that recognition of the disciplinary obligation in the PROSE instrument emphasizes accurate or elegant representation of the mathematics (e.g., correctness) over engagement with deeper content. Thus giving more challenging work would lessen the chances of students producing mathematically accurate work. Adding items to the PROSE-disciplinary instrument that address issues of engaging students with authentic instances of doing mathematics might yield different results.

In general, I find it notable that professional obligations tend to enable rather than constrain inquiry-oriented practices thus far. I had hypothesized that they would explain inconsistencies between beliefs and practice, and there may be ways of differing or creating additional items for the PROSE instrument that would help model instructors' true difficulties in implementing inquiry-oriented practices.

Student-Teacher Models

The use of *interactive lecture* had three significant indicators: (1) instructors' belief on *modeling* for incremental mastery predicted less use of practices of interactive lecture (ß=-.65, $p<.05$), (2) instructors' *awareness* of students' mathematical dispositions predicted more use of practices of interactive lecture (ß=.54, $p<.01$), and (3) recognition of the interpersonal obligation predicted less use of practices of interactive lecture (ß=-.42, $p<.05$), shown in Table 6.9. It seems reasonable that an instructor who believes they should model things for their students by directly showing them how to do it would not want their lectures to be as interactive. This is notably the only inquiry-oriented practice for which recognition of the interpersonal obligation is a significant predictor. It could help explain why an instructor whose beliefs predicted frequent use of interactive lecture still does not use it a lot. I expected that the interpersonal obligation was negatively related to the practice of making lecture more interactive either because (1) an instructor concerned with the interpersonal or social environment would actually try to minimize use of lecture altogether in favor of other formats, like working in groups or (2) an instructor concerned about the interpersonal environment in terms of making things fair and equitable could be worried that more interactions could lead to inequitable interactions. Thus, the results shown by the significant relationships between beliefs, professional obligations and the practice of interactive lecture confirm the hypothesis that professional obligations can help explain a misalignment between beliefs and a particular inquiry-oriented practice.

Table 6.9: Results from structural models predicting student-teacher latent constructs, n=269

		Interactive Lecture		**Hinting and not telling**	
Instrument	Predictor	Standardized Estimate	SE	Standardized Estimate	SE
Beliefs	Struggle	-.10	.20	.60**	.18
	Model	-.65**	.22	.32	.21
	Awareness	.54***	.11	.08	.11
Professional Obligations	Individual	-.09	.16	.07	.16
	Interpersonal	-.42*	.19	.13	.18
	Disciplinary	.36†	.20	-.15	.18
	Institution	.14	.14	.16	.12

† $p<0.10$, *$p\leq0.05$, **$p\leq.01$, ***$p\leq.001$

Instructors' decisions to approach responding to students by *hinting without telling* was predicted by the belief that students should be allowed to struggle (ß=.60, $p<.01$). Other environmental factors, as represented by the professional obligations, did not play a significant role in predicting this practice. The practice of hinting without telling is not as distinct from more typical forms of instruction, that is, it can be done within a lecture format without changing virtually anything in the lesson plan or curriculum. This might explain why recognition of various professional obligations had no significant impact on the decision to do this practice.

Student-Student Models

The student-student inquiry-oriented practices of including group work and student presentations were significantly predicted by both beliefs and professional obligations (see Table 6.10). The practice of choosing to have students participate in *group work* with their peers was significantly predicted by the instructors' belief that students should be allowed to struggle (ß=.46, p<.01), awareness of students' mathematical dispositions (ß=.27, p<.01), recognition of the individual obligation (ß=.28, p<.05) and recognition of the institutional obligation (ß=.28, p<.05), as shown in Table 6.10. The tendency to attend to individual students' needs seems necessary in order to do the extra work of attending to students' individual problems and struggles as they work in groups. During other class formats such as lecture, individual students' needs are only prevalent when students choose to raise their hands, whereas during group work an instructor is virtually obligating themselves to look for ways to attend to individual students' needs. I am less clear how the institutional obligation significantly predicts the use of group work. Though there is no evidence to support that mathematics departments require the use of group work in their lower-division courses, I expect that this is linked to the widespread pressures for mathematics departments to include more active learning and student engagement, as outlined in the theoretical framework. I had expected recognition of the interpersonal obligation to be a positive significant predictor of group work, because I thought that someone who recognized the need of fostering the social environment of their classroom would be more likely to use group work. On the other hand, I suspect that, to many instructors, fostering a productive social environment might not have anything to do with using group work or not, but rather *how* the instructor manages group work or a whole-class discussion.

Table 6.10: Results from structural models predicting student-student latent constructs, n=260

		Group work		Presentations	
Instrument	Predictor	Standardized Estimate	SE	Standardized Estimate	SE
Beliefs	Struggle	.46**	.17	.63***	.18
	Model	-.04	.19	.31	.20
	Awareness	.27**	.10	.15	.11
Professional Obligations	Individual	.28*	.14	.64***	.14
	Interpersonal	-.01	.16	.04	.17
	Disciplinary	-.13	.17	-.53**	.18
	Institution	.28*	.11	.21†	.12

† $p<0.10$, *$p\leq 0.05$, **$p\leq .01$, ***$p\leq .001$

The practice of having students give presentations was significantly predicted by the instructors' belief that students should be allowed to struggle (ß=.63, $p<.001$), the instructors' recognition of the individual obligation (ß=.64, $p<.001$), and the instructors' recognition of the disciplinary obligation (ß=-.53, $p<.01$). Even if an instructor wanted his or her students to struggle with the mathematics, they might not mandate student presentations in an effort to maintain control over the representations of mathematics. Then the instructor can make sure the mathematics being presented is accurate, clear and elegant, as Mahavier (1999) suggested. On the other hand, if an instructor feels strongly about attending to individual students' needs, they might be more inclined to have students give presentations for each individual students' benefit and have the empathy or patience to wait for students to figure things out themselves. Thus, the results shown by the significant negative relationships between the belief that students should be allowed to struggle and the recognition of the disciplinary obligation predicting the practice of having students present confirm the hypothesis that professional obligations can help explain a misalignment between beliefs and a particular inquiry-oriented practice. Additionally, the individual obligation and the institutional obligation show that, in some instances, professional obligations can also enable the use of inquiry-oriented practices.

All Models with Partial Sample A for Comparison with Total Sample Results

As in the previous chapter, in addition to presenting results with all data that were collected (the total sample), I also present results without the participants who claimed to never ask students to revise definitions and did not retake the survey (partial sample A), in an effort to demonstrate the potential impact of the missing data.

The models run with listwise deletion of missing data (e.g., without data from the participants who did not complete the INQUIRE instrument due to the question about definitions) tended to be similar but with more extreme patterns, as shown in Table 6.11, Table 6.12, and Table 6.13. The data included in the models in the previous tables, Table 6.8, Table 6.9, and Table 6.10, include participants who did not complete all the INQUIRE items, meaning that their data would make the relationship between inquiry-oriented practices and beliefs and professional obligations weaker. While contributing to the normalization of the beliefs and obligation scores, the additional participants in the total sample (as compared to the partial sample) do not contribute to the corresponding INQUIRE data. Thus, the patterns appear more pronounced among the sample (partial sample A) that contains only the people that completed all instruments.

First comparing the student-content models, in the model with the total sample, I found that instructors' belief that students should struggle was positively associated with all three practices, but not always significantly. Then, in the models with partial sample A, instructors' belief that students should struggle was a positive significant predictor for all three practices. That is, the coefficients are greater and the *p*-values are smaller for each of the three incomplete models in Table 6.11, compared to the coefficients and the *p*-values in Table 6.8. Recognition of the interpersonal obligation is significant in all the incomplete models, and is not in the models with all participants. The latent constructs that had smaller *p*-values predicting outcomes in the original models than in the incomplete models were instructors' awareness of students' dispositions and recognition of the institution obligation predicting the practice of providing students with open problems, and instructors' awareness of students' dispositions and recognition of the individual obligation predicting the practice of giving students opportunities to construct and critique.

Table 6.11: Models predicting student-content latent constructs using partial sample A data, n=194/5

		Construct/Critique		**Prove**		**Open Problems**	
Instrument	Predictor	Standardized Estimate		Standardized Estimate	SE	Standardized Estimate	SE
Beliefs	Struggle	.49**	.19	.44*	.21	.79***	.22
	Model	.16	.21	.12	.22	.14	.24
	Awareness	.27*	.12	.28*	.13	.19	.14
Professional	Individual	.26†	.15	-.09	.17	.46*	.18
Obligations	Interpersonal	-.40*	.20	-.53*	.23	-.49*	.03
	Disciplinary	.03	.20	.31	.24	.07	.76
	Institution	.03	.12	-.04	.13	.30*	.14

† $p<0.10$, *$p\leq0.05$, **$p\leq.01$, ***$p\leq.001$

Table 6.12: Models predicting the student-teacher latent constructs using partial sample A data, n=194

Instrument	Predictor	**Interactive Lecture** Standardized Estimate	SE	**Hinting and not telling** Standardized Estimate	SE
Beliefs	Struggle	-.10	.26	.71***	.22
	Model	-.57*	.28	.40	.26
	Awareness	.47**	.16	.07	.14
Professional Obligations	Individual	-.17	.20	.18	.18
	Interpersonal	.53†	.28	.10	.24
	Disciplinary	.57*	.27	-.16	.23
	Institution	.11	.17	.25†	.14

p<0.10, **p*≤0.05, ***p*≤.01, ****p*≤.001zq

Table 6.13: Models predicting the student-student latent constructs using partial sample A data, n=194

Instrument	Predictor	**Group work** Standardized Estimate	SE	**Presentations** Standardized Estimate	SE
Beliefs	Struggle	.50**	.18	.70***	.21
	Model	-.03	.20	.39†	.24
	Awareness	.23**	.11	.06	.13
Professional Obligations	Individual	.29*	.15	.67***	.16
	Interpersonal	-.04	.19	.05	.21
	Disciplinary	-.10	.19	-.53*	.21
	Institution	.30**	.12	.22†	.13

† *p*<0.10, **p*≤0.05, ***p*≤.01, ****p*≤.001

Comparing the student-teacher models in Table 6.9 to models in Table 6.12, there is a similar pattern with the results of the partial sample being slightly more predictive than in the model with all participants. That is, the significant predictors of the practice of interactive lecture in Table 6.9, instructors' belief in modeling for incremental mastery, and instructors' awareness of students' dispositions, are also significant in Table 6.12. Additionally, the recognition of the disciplinary obligation is a significant predictor in Table 6.12, but not the recognition of the interpersonal obligation. The sole predictor of the practice of hinting and not telling is the instructors' belief that students should be allowed to struggle. This last predictor is significant with a smaller *p*-value in the model using the partial sample. Finally, comparing the student-student models in Table 6.10 and Table 6.13, there are almost identical results. The only distinctions, in terms of significance, were that the recognition of the institutional obligation is significant with more certainly as a predictor for the practice of group work in Table 6.13, and the recognition of the disciplinary obligation is significant with more certainty as a predictor for the practice of having students give presentations in Table 6.10. Thus, the claims made in the previous sections are likely valid, even considering the participants that did not return to complete the INQUIRE instrument. I turn to discussing the implications of these results.

Chapter 7 **Discussion and Conclusion**

In this closing chapter, I first discuss the theoretical contributions of this work in relation to the research questions. I then move to possible implications for policy makers, both in terms of understanding the contextual environments of the instructors they are trying to influence, and how advocates of IOI might leverage results from this study to strengthen their recommendations. Then I highlight the contributions from the instruments created for this study, the methods used to analyze them, and modifications recommended for future iterations. I comment on future plans to use the INQUIRE instrument in conjunction with other instruments, particularly those with an equity focus. Finally, I mention limitations of this work and make closing remarks.

Theoretical Contributions

This section is organized by research questions, which were as follows: (1) what inquiry-oriented instructional practices are used by college mathematics instructors and, (2) what factors might predict their use? First, I explain what inquiry-oriented instructional practices are used by college mathematics instructors, and explain the impact of this framework for our knowledge about IOI and the norms of the didactical contract that are breached by attempting any one of the inquiry-oriented practices. I compare the findings to another theoretical framing of IOI. Then, I turn to theoretical implications of the findings relevant to the second research question. That is, what the results of this study can contribute to the knowledge about teacher decision-making, using the framework of practical rationality.

Relevant to the first research question, the existing literature on inquiry often referred to IBL or IOI as though everyone had a shared understanding of what it meant and that it was a single construct, when researchers and practitioners were sometimes referring to different components of inquiry. Though there have been efforts to consolidate conceptions of IOI through frameworks like Laursen and Rasmussen's pillars (2019), no effort had been made to systematically, empirically parse the components of IOI and see where they are found together. This study fills the gap in the literature on inquiry by organizing inquiry-oriented practices in the literature using the instructional triangle. That organization was then empirically tested by the INQUIRE instrument, which led to the identification of seven practices that instructors use in their classrooms with wide variation. They include the practice of including open problems, providing opportunities to construct and critique conjectures or claims, construct formal arguments, making lecture interactive, hinting without telling, organizing the class in a group work format, and including student presentations.

These seven practices give insight to the ways that instructors attempting to incorporate inquiry must breach the didactical contract. I hypothesized that instructors would potentially be breaching division of labor, temporal, and exchange norms; these practices offer evidence of all three. For the norms concerning division of labor, constructing claims or conjectures, constructing formal arguments, telling students how to proceed, and presenting new material to the class is often the work of the instructor. Incorporating all seven of these practices requires the instructor to breach those norms by placing some of the work on the student. In terms of temporal norms, all of the practices would often require a different amount of time to convey the knowledge at stake. For example, if a student is struggling and the instructor gives a hint rather than telling the student exactly what to do next, it would likely take the student longer to complete the problem. Some versions of IOI might require a different temporal sequence of events. Instead of first hearing about a topic from the instructor, seeing an example on the board, and then working on it alone, the order might be reversed. So in a class emphasizing student engagement with the content, the student might first work alone on a problem, then see some classmates do an example on the board, and then hear the instructor give a short lecture to debrief on the work done. In terms of exchange norms – what needs to happen and how – constructing and critiquing conjectures or claims and including open problems are breaching norms concerning what needs to happen in order to communicate the knowledge at stake. For example, in a traditional class, perhaps what would "count" as communicating the knowledge at stake would be explaining the topic with some visual aids on the chalkboard. This study, on the other hand, shows that if an instructor is trying to incorporate the practice of interactive lecture, then the norm would need to be breached so that the knowledge at stake would not be considered communicated until students had a chance to try some examples on their own or demonstrate they were following the instructor's explanation. A key difference in the instructional exchanges using a few practices, including making lecture interactive, hinting without telling, and group work, is that the transaction of mathematical knowledge is expected to be constructed jointly through interaction, rather than one party (be it the instructor or students) telling everything in a undirectional way. Aside from developing the framework that resulted in the naming of seven practices, this study also looked at how those practices tend to be used together.

The cluster analysis showed that these practices break into roughly four different patterns of use by instructors, three of which are characterized by practices recommended by those who promote inquiry-oriented learning. This result, in particular, supports the argument that IOI is not a single practice that distinguishes instructors among those that do and those that do not practice it. Rather, the practices recommended by promoters of IOI support at least three distinct clusters of instructors (in comparison to the fourth which includes people who use all practices less frequently), distinguished by different sets of practices, all of which are desirable. This finding has profound implications for future research and possible suggestions for those that engage in policy making.

In the introductory chapter, I posed that this study would bridge work done in the RUME community, on teaching with inquiry, and teacher decision-making. The identification of seven practices organized by the instructional triangle creates a framework on IOI that draws from the inquiry literature, contributes to education researchers' knowledge of the didactical contract, and is applicable to the college mathematics context. Conversely, the framework developed for the college mathematics context can be leveraged for other contexts where inquiry is applicable. The framework that categorized inquiry-oriented practices could then be used alongside theories of decision-making to understand why some instructors use those practices or collections of practices more frequently than others. As a check to the validity of these findings, I now use my findings and work backwards to see if they align with conceptions in the literature.

An Exercise in the Validity of the INQUIRE framework

The four pillars proposed by Laursen and Rasmussen (2019) offer an outside frame of reference for of the characterization of IOI I am offering given the empirical results of the INQUIRE instrument. I chose to compare the INQUIRE results against their conception because it was created from a merging of both IOI and IBL, and based on both researcher and practitioner accounts. They proposed four pillars of inquiry-based mathematics education including: (1) that students engage in meaningful ways with mathematical tasks, (2) that students collaborate to make mathematical meaning, (3) that instructors elicit and value students' ideas and (4) foster equity. The practices from the INQUIRE instrument that emphasize student interaction with the content, that is, giving students the opportunity to construct and critique claims, having them construct formal arguments, and work on open problems, all map to Laursen and Rasmussen's (2019) first pillar of having students "engage deeply" with the mathematical content. The practices involving interactions among students, including group work and student presentations to each other, both map to the second pillar of student collaboration. The practices that emphasize instructor interactions with the student, including making lecture more interactive and giving hints without telling students what to do, both map onto the third pillar of valuing and making space for student thinking.

In addition to being able to map from the INQUIRE instrument back to the pillars, the pillars themselves suggest potential areas for improvement of the INQUIRE instrument. For example, drawing from the first pillar, one could ask whether there are ways that students can engage meaningfully with mathematical tasks that are not currently covered in the INQUIRE instrument. In the section on future INQUIRE iterations, I explain the implications of the current lack of practices in the instrument that map on to the fourth pillar.

This mapping from the INQUIRE's framework to the four pillars shows that the organization of the INQUIRE instrument not only is covering all interactions by its design around the instructional triangle, but is additionally justified on account of theory developed from a partnership of past IOI and IBL research. The cluster analysis in this study shows how these pillars could be realized in classroom practice in different ways. Two instructors could try to enact the first pillar equally, but one from the first cluster could enact it by giving students more open problems, while one from the second cluster could enact it by giving students more opportunities to construct and critique claims, and create proofs. Given the theoretical soundness of the INQUIRE instrument and findings, the natural question that follows is why or how instructors have chosen to use those inquiry-oriented practices.

Practical Rationality Framing the Use of Inquiry-Oriented Practices

The second question addressed what factors might predict the use of inquiry-oriented practices by bringing in the theory of practical rationality to consider not only individual but also social factors. This work frames inquiry-oriented practices as breaches of the didactical contract instructors and students are usually familiar with in the undergraduate mathematics setting that require negotiating, often led by the instructor. Because it is not simple to undertake a renegotiation of the exchange of mathematical content, it follows that instructors might be motivated to do so due to beliefs that inquiry-oriented practices would allow their students to learn better. Professional obligations could further enable or constrain those efforts. Thus, the second research question was answered by responding to two specific research questions:

a. Is there a relationship between the beliefs of instructors and inquiry-oriented practices or patterns of inquiry-oriented practices in college mathematics instruction?
a. If so, do professional obligations help explain the relationship between beliefs and inquiry-oriented practices and patterns of inquiry-oriented practices in college mathematics instruction? If so, how?

The answers to these questions were affirmative. As hypothesized given the framework of practical rationality, the individual predictors of instructors' beliefs and social predictors of professional obligation recognition significantly predicted each separate practice and pattern of practices. What was not expected was the way in which different professional obligations would positively or negatively predict the particular inquiry-oriented practices.

The results showed that professional obligations helped explain discrepancies between beliefs and practice. Instructors with *learner-focused* beliefs (e.g., high belief that students should be allowed to struggle, low belief that modeling is the best for incremental mastery) were significantly more likely to implement almost every inquiry-oriented practice. However, the mean use of some of the practices was better explained once recognition of professional obligations were added to the models – consistent with the literature that found beliefs do not always align with practice. The practices of including more *interactive lecture* and *student presentations* had as significant negative predictors of recognition of the *interpersonal obligation* and the *disciplinary obligation,* respectively, indicating that those professional obligations were in direct opposition to the learner-focused beliefs. Instructors' individual beliefs might give them reason to deviate from the norms of teaching mathematics in undergraduate classrooms, but recognizing some aspects of the obligations inherent to their jobs could steer them back to the norm. These results depend partly on how the professional obligations are represented in the PROSE instrument.

I acknowledge a caveat to these results: what constitutes each professional obligation in the results is dependent on what items were retained in the PROSE instrument. So, when I say that the recognition of the disciplinary obligation is a negative predictor of instructors' reported use of student presentations, I cannot claim that recognition of any interpretation of the disciplinary obligation would lead to less inquiry. If the PROSE-disciplinary instrument was expanded to include more items that emphasized the importance of having authentic experiences struggling with the mathematics as mathematicians would encounter in their regular work, rather than focusing on items that emphasized the importance of accuracy and elegant solutions, I would expect the recognition of the disciplinary obligation would be positively related to frequency of student presentations. The takeaways from this caveat are that the professional obligations are multifaceted and the results should be interpreted as such, and that there is more work to be done to develop the PROSE instrument and understand what are the substantive interpretations of the professional obligations for instructors. Future work could investigate whether recognition of the interpersonal and disciplinary obligations always functioned as negative predictors of inquiry-oriented practices, or if different representations of the obligation could also enable inquiry-oriented practice.

Conversely, the social resource of professional obligations also helped explain why instructors implemented some of the inquiry-oriented practices more often than their beliefs would indicate. The inquiry-oriented practices of *group work, student presentations,* and including opportunities to *construct or critique* and solve *open problems* were significantly predicted by the recognition of the *individual and institutional obligations.* The membership in cluster with the overall highest use of inquiry-oriented practices was also predicted by cognition of the individual and institutional obligations. I had hypothesized that recognition of the institutional obligation would act more as a constraint and predict more normative instruction, but instead it was related to using inquiry-oriented practices. While I did not ask participants if their department requires or recommends the use of IOI or IBL, the majority of participants reported in the background survey that they were not using them. However, a lot of participants are using inquiry-oriented practices without labeling it with that name. This could be due to institutional pressures to include more active learning and student engagement, which was translated by mathematics departments and instructors to mean more novel tasks (i.e., open problems) and group work. The theoretical implication from these results is that both beliefs and professional obligations have the potential to pull instructors in both directions – towards complying with the didactical contract or strategically breaching it, depending on the practice and obligation at stake.

A closely related finding was that while beliefs and professional obligations are predictive of each practice, they do so in different ways for each practice. That is, the framework of practical rationality applies to each practice, but in unique ways to each. These findings provide nuance to the popular idea that instructors either teach with IOI or they do not. With its diverse history from both K-12 research and undergraduate instruction, IOI has had many diverse meanings. This links back to the importance of having an organized breakdown of what is meant by IOI and what components of it are used by instructors. The literature on inquiry and on decision-making can thus symbiotically benefit from this study that utilizes both. Beyond a theoretical perspective, these results could also be useful for advocates or potential practitioners of IOI.

Practical Implications

When policymakers offer research-based recommendations, they often do not consider the complex contexts that instructors operate within. I introduced this study by framing it with the growing need for innovative teaching methods and the documented success of active instruction. Despite well-intended recommendations, there are many instructors who frequently instruct without active components. Mahavier (1999) conjectured that instructors do not use IOI due to impatience, and others might claim that instructors are simply lazy or too stubborn to change their habits. This study shows that beliefs about instruction often do not align with executed practice because there are other environmental and social factors that provide reasonable constraints on their range of realistic options, as modeled here by professional obligations. Instructors are left to sort out conflicting demands, leading to the internal struggles mentioned by

Lampert (1985). For example, even if instructors are convinced that having more student presentations would create the most learning opportunities, they might decide to lead more whole-class discussions for the sake of ensuring that the mathematics the students are exposed to is shown in an accurate and elegant representation. Having more interactive lecture might align with an instructor's desire to model less, but not enough to risk that students could be at risk to have less equitable participation. So even if policy makers give convincing recommendations that instructors theoretically agree with, there are still rational reasons for not altering their instruction. This study provides motivation for more work like Yoshinobu and Jones' (2012) *The Coverage Issue* that addressed the institutional obligation by way of the fear of instructors that they would not be able to cover as much material using inquiry-based learning than other methods of instruction.

Recognizing that the disciplinary obligation and interpersonal obligation play a role in the decision to use IOI gives policymakers the opportunity to address them directly, if they want to advocate for the implementation of inquiry-oriented practices. For example, if instructors are worried about mathematics being represented accurately by students, policymakers could offer suggestions on how to scaffold student presentations in a way that would allow them to get input from others before putting up the final product for the whole class. Or policymakers could work with researchers to investigate the impact of putting up inaccurate work (e,g., Do other students retain the inaccurate information? Are they able to notice what is not accurate? Does it help student affect or learning to see and discuss unfinished work in progress?).

Conversely, the results from this study give two ways of triangulating which practices might be more feasible to implement. First, the descriptive statistics from the INQUIRE instrument shown in Table 6.1 and Table 6.2 show which practices are avowedly being implemented more frequently. The practices of interactive lecture and hinting without telling have higher means than the other practices, indicating that instructors are already choosing to use those components of active instruction. Second, recognition of professional obligations was a positive predictor of the practices of giving students opportunities to construct and critique claims, solve open problems, and participate in group work. Recognition of professional obligations was not a significant negative predictor of the practices of including more opportunities to prove, and hinting without telling. If policymakers, departments, or universities are attempting to advocate for more active learning practices, these results highlight which practices instructors would be more inclined to act on. In the instances where recognition of professional obligations are positive predictors, the obligations could be leveraged to advocate for particular inquiry-oriented practices. For example, if instructors are not already having students construct and critique claims during class time, appealing to instructors' desire to meet individual students' needs might be one strategy to encourage this practice.

The results of this study have possible implications for the types of professional development that might accompany an initiative to include more IOI. The recognition of professional obligations had the largest effect sizes on the least used practice (i.e., student presentations). For the practices of interactive lecture and student presentations, which had significant negative associations with the recognition of the interpersonal and disciplinary obligations, respectively, professional development can adapt to the challenges identified by the results. For the practices that are positively associated with recognition of professional obligations, the professional obligations can be leveraged to encourage those practices. For example, if a department wanted instructors to include more opportunities for students to solve open problems, they might appeal to instructors' recognition of the institutional obligation by making it part of coordinated homework assignments for the course. Or they could hold workshops that allow instructors to practice attending to students as they solve open problems. Though these practical implications are drawn from relating the results from the PROSE instrument and the INQUIRE instrument, the PROSE instruments now exist and could be used for other purposes.

Methodological Contributions

This study offers two new instruments, a sampling strategy for studies concerning IOI, and a methodology (SEM) for operationalizing a conceptual framework for understanding teacher decisions in IOI classrooms. The psychometric work presented in Chapter 5 to confirm the grouping of items onto factors helps to advance the theory behind what it means to implement IOI and how to measure the professional obligations of college mathematics instructors. The creation, development, and validation of the INQUIRE instrument opens a path for future work.

The INQUIRE instrument parses out components of inquiry-oriented practice and patterns of practices for quantifiable, large-scale study. As outlined earlier, general inquiry-oriented practices have shown mixed results considering issues of equity. For example, Johnson and colleagues (2018) found that IOI did not increase equity for the students in their study. In fact, it worsened some achievement gaps. They conjectured that the reason that their study had different results than Laursen and colleagues' (2014) report was because the courses in Johnson's study heavily relied on group work, while Laursen's relied on student presentations. Johnson and colleagues (2018) wondered if equal opportunities could be controlled better by choosing students to present rather than leaving students in more control of their own participation through group work. The INQUIRE instrument could be used to empirically investigate such hypotheses. Whether or not the INQUIRE instrument is used, studies concerning IOI could make use of the cluster analysis in this study.

The cluster analysis from the INQUIRE instrument revealed that instructors fall into one of three patterns of inquiry-oriented instruction. The first included instructors who reported using all the IOI practices with high frequency, the second especially reported emphasizing practices that had students working directly with rich mathematical content, and the third reported lower use of most of the practices except interactive lecture and hinting without telling. For future studies, these clusters offer a sampling strategy. For example, if a researcher wanted to do classroom observations of IOI, these three groups and the quantity of instructors I found in each could offer a guide to choosing a proportionally balanced sample of the variety of patterns of IOI instruction. There could also be work to classify types of IOI used in accounts of existing research, in case there could be similarities within outcomes. For example, the equity work done by Laursen et al. (2014) might have included more instructors in the first and third clusters (as they included student presentations) while the work by Johnson et al. (2019) might have focused more on instructional practices emphasized by the first and second clusters. This might explain why the implications for equity were contradictory. By retrospectively classifying studies by the type of instruction, the outcomes might give a more consistent story.

The college PROSE scenario-based instrument was the second instrument developed for this dissertation, having been originally designed for K-12 teachers. Its adaptation and initial psychometric work with college instructors give evidence that professional obligations are applicable to the college-level instructor population. The factor analyses, based on patterns in the participants' responses, give insight to what the professional obligations are composed of. Professional obligations are relevant to any instructional decision. Other mathematics education researchers can choose any instructional issue of interest (e.g., hybrid courses, noticing, responding to students, etc.) and administer the PROSE instrument in combination with gathering data on the issue of interest. The PROSE data would give information about how the instructors' recognition of professional obligations plays a role in relation to how they handle any of those particular issues.

Apart from the PROSE instrument specifically, scenario-based assessments are not widely utilized in the mathematics education community. This study shows their usefulness for attaining responses from a large-scale population on higher-inference constructs. I would question the validity of self-report responses had I asked survey questions like, "Do you value attending to the needs of your individual students?" due to participants trying to please the researcher. The PROSE instrument used scenarios to investigate that question in a way that reflected what instructors might actually do, as opposed to what they are expected to answer. Oftentimes researchers worry that the only way to collect such data is through video observations, but this type of instrument offers another method to do so. I did not use scenario-based assessments for the INQUIRE instrument because I asked questions about frequency of concrete actions rather than more general opinions on something. That is, questions like "If a student asks you to look at his or her work during class, how often do you respond without evaluating whether or not it was correct?" were straightforward to answer. But had I provided various depictions of an instructor not evaluating whether students work was correct, and asked how frequently participants did what was depicted, they might pick up on unintended details. But scenario-based assessments are useful to study the use of high-inference aspects of IOI. For example, I would not trust respondents to accurately respond to a question like, "How often do you try to make student thinking visible?" but would rather design a wide variety of scenarios where participants could choose to make student thinking visible or not, and see what they choose. This study not only created new instruments, it also linked them together using a method not yet widely used in the field of mathematics education.

Structural equation modeling (SEM) was uniquely useful to link all the instruments together and test the theorized relationships between them. Unlike OLS regression where relationships must be tested in sequential steps, SEM allows the simultaneous measurement of all relationships, both item (or parcel, in this case) to construct, and constructs with other constructs. In an OLS regression, there would be one dependent variable predicted by independent variables, but with SEM there can be multiple dependent variables and their relationships to each other can be examined. My models were relatively simple and straightforward due to the nature of the research questions and size of the sample. With a larger sample, future models could include all seven regressions from the results section simultaneously. Using SEM opens the window for testing more complex relationships. For instance, if a researcher was interested in how inquiry-oriented practices interact with each other, (e.g., does more group work imply that students are given more open problems to work on together?), they could test paths between them, while simultaneously keeping all the independent variables (beliefs and obligations in this study) in the model. While SEM is fully embraced by fields such as psychology, it has been slower on the uptake by mathematics educators.

Future Directions

There are two ways I aim to develop the INQUIRE instrument aside from the content improvements noted in the scale development chapter: First, I plan to investigate the predictive validity of the INQUIRE instrument by developing a companion classroom observation instrument. The work to translate the survey to an observation instrument would involve rephrasing the questions and creating a protocol for coders. In the survey, questions are of the form, “How often do you [insert inquiry-oriented practice here]?” and respondents can select options from a Likert-type scale ranging from Never to Multiple times per class. The observation instrument would rephrase the questions so that each inquiry-oriented practice would be listed, and coders would select if that practice was used in a given unit of class time. Once the instrument is developed and piloted, it would be filled out as a log, independently by researchers and the instructor, to assess the predictive validity of the self-reports.

Second, I aim to expand the use of the INQUIRE instrument in conjunction with more instruments measuring constructs beyond beliefs and professional obligations. A first potential new lens could involve a different framework for decision-making. In this study, I focus only on conscious, rationalized decision-making. I would argue, however, that many decisions in the classroom are made without much conscious thought. Instructors make in-the-moment decisions throughout a single class, and though they can be rationalized after the fact using a theory like practical rationality, it does not explain what is occurring in their thought processes to bring about the given result. The literature on decision-making includes well-developed notions like intuition, which are decisions made off rapid, and non-rational unconscious judgements (Dane & Pratt, 2007). Unlike other types of decisions that are made based on conscious justifications, these decisions are made based on holistic views and associations of patterns that experts acquire over time (Dane & Pratt, 2007). In the classroom, I imagine instructors would develop intuition for how to follow the didactic contract unique to their context as they gain experience. Research on how to disrupt the intuition to follow normative practice or ways to retrain some habits of mind in the context of developing new norms for instructional exchanges would give additional insight to the use of inquiry-oriented practices.

A second potential new lens could include student perspectives, gauging the impact of each individual practice on students' experiences. If instructors choose not to use student presentations due to their recognition of the individual obligation, that is likely to be the case based on their own perceptions of students' experiences. But what if students' experiences were different than instructors think, and what if instructors had access to data about how student presentations impact their student experience? Would that make a difference in their use of this practice? In the literature review, I outlined the importance of IOI for equitable instruction as a motivation for knowing more about whether instructors decide to use IOI. However, this study did not investigate how IOI practices influence equity, or its lack thereof, in student experiences. In popular images of IBL, students have more opportunities to interact with each other and the mathematics, which can give more opportunities for inequity to surface. The increased interactions between teachers and students, and students with each other, can give rise to many more opportunities to exacerbate or reduce inequities than with a lecture format (where inequities certainly exist but are less visible). Group work in classrooms has been identified as a "locus for negotiation of power and cultural norms" (Takeuchi & Bryan, 2018, p. 124). As an example of how power can influence outwardly objective mathematics interactions, students can ascribe mathematical authority based on status rather than reasoning. I noted earlier how even professional mathematicians sometimes rely on other indicators (such as the status of other mathematicians who reviewed the proof) rather than mathematical content to judge the soundness of published proofs (Weber & Mejía-Ramos, 2011). Data on how inquiry-oriented practices impact individual or groups of students' affective experiences, e.g., if they feel a sense of belonging, if they experience microaggressions, if they feel confident, if they feel motivated, etc. could inform instructors' decisions and departmental recommendations.

Limitations

A limitation of the survey data collected is the sample, both in terms of the method that it was attained and its size. It would have been difficult to randomly sample from all college mathematics instructors around the U.S. There might be response bias built into my data, as instructors that put more effort into their teaching practice might be more inclined to invest the time to reflect on it. This might skew results about the prevalence of IOI. To prevent this, I attempted not to advertise that I was more interested in innovative or inquiry-oriented instructional practices than any other pedagogies or advertise what the study was about (e.g., the name of the instrument was not seen by the participants). This response bias would not have affected the validity of the relationships between constructs in the model. The other issues of self-reported data are considered and addressed in the methods section.

In terms of sample size, I would have liked to create SEMs for this study that involved more or all of the instruments simultaneously in the same models. This could give a better sense of the relationship between practices. However, I did not have enough participants for the number of free parameters that those models would have required. A model with one practice and the beliefs and obligations contains 103 free parameters, and a model with seven practices with beliefs and obligations would require 220 free parameters, quickly approaching the total number of participants.

A second limitation is derived from the nature of this study, how it separates IOI, beliefs, and professional obligations, in order to investigate how they interact. In a study of students learning to write proofs and experiencing mathematically satisfying moments, Satyam (2018) sensed that the process of breaking complex phenomena down and recomposing its components lost something that was going on. She found that the whole was greater than the sum of its parts, and her methodology evolved to study the interactions between proof writing and satisfying moments as a interactive whole, rather than two separate components. This more holistic approach had a long history grounded in qualitative work like phenomenological studies in education, which focus on the meaning of phenomena (such as IOI) for those experiencing it rather than on the objective reality of what something is (Henriksson & Friesen, 2012). Similar to Satyam's original strategy, this study separates components and then reconnects them. That is, I break down the phenomena of deciding to use IOI into a lot of pieces – seven inquiry-oriented practices, three beliefs, and four professional obligations. While I argue that the work done here is valuable and fills a gap in the literature, I also recognize that other studies that treat IOI as a whole could capture things that this study cannot.

Conclusion

Inquiry-oriented instruction is not a homogenous practice. In the literature, researchers and practitioners often refer to IOI or IBL and use a wide variety of differing definitions. This study offers evidence that IOI can take different forms, and information about what those forms tend to look like in practice. Future research can benefit from explicitly doing work with reference to one or more specified versions of IOI. This study offers a theoretical conception with empirical evidence of why instructors choose to use some patterns of inquiry-oriented practices over others.

Whether instructors choose to use inquiry-oriented practices or not, this study offers evidence that those choices are rational. An examination of how they make decisions before questioning their practice can help inform how to partner with instructors instead of change them. The framework of practical rationality helps explain why beliefs do not always align with the implementation of inquiry-oriented practices, because recognition of the interpersonal and disciplinary obligations can be in direct contradiction with learner-focused beliefs. I gave practical implications for those wishing to influence practice: First, findings from this study suggest that some practices as more feasible and thus might make better gateway recommendations for advocates of IOI. And second, working with instructors to problem-solve about how they can carry out their beliefs and simultaneously fulfill their professional obligations could be useful to instructors so they are not left to struggle with the inner turmoil of not being able to realize their beliefs about student learning on their own. A direction for future research would be to investigate how different components or patterns of inquiry-orienting instruction benefit or harm various student populations, and what instructors can do to make the best experience for all students.

Appendices

Appendix A: INQUIRE instrument

The following items are randomized, and instructors can respond on a 6-point Likert-type scale with the following options: 1-Never, 2-A few classes, 3-Less than half the classes, 4-More than half the classes, 5-Every or almost every class, or 6-Multiple times per class. The survey is administered twice – once for a lower-division class and once for an upper-division class. If an instructor has not taught such as course, Qualtrics does not administer the survey. The titles (e.g., student-content interaction) are not shown to participants.

Welcome to the survey on instructor orientation. Each question asks you to consider how often you do something in your typical teaching practice.

Lower-division prelude: What course will you be thinking of as you continue? Please do not choose a course taught completely online.
Choices: Calculus I, II, III, or IV; Introductory Linear Algebra, Differential Equations, AN introductory course such as college algebra, pre-calculus or trigonometry, Other, I have not taught a lower-division course.

Upper-division prelude: What course will you be thinking of as you continue? Please do not choose a course taught completely online.

Choices: Modern Algebra, Analysis, Topology, Advanced Linear Algebra, Combinatorics, (Topics in Mathematics) for Future Teachers, Introduction to Proof, Other, I have not taught an upper-division course

To answer these questions, recall what you did during the most recent semester or quarter you taught a typical class. Do not consider homework, exams, office hours, or other interactions outside of class.

Table A.1: Items relevant to student-content interactions from the INQUIRE instrument and their factor structure from an EFA

Item*	1	2	3	Statement
lscdef_1	0.668*	0.075	0.061	How often do you ask students to revise a definition?
lscdef_2	0.730*	0.066	-0.016	How often do you ask students to propose a definition?
lscdef_3	0.723*	0.058	0.059	How often do you ask students to critique a proposed definition?
lscdef_4	-0.078	0.824*	0.061	How often do you ask students to figure out how to use a definition?
lscdef_5	0.002	0.780*	0.2	How often do you ask students to figure out when to apply a definition?
lsclak_1	0.656*	0.138	0.038	How often do you ask students to generalize a claim?
lsclak_2	0.696*	0.128	0.017	How often do students propose a hypothesis for a theorem?
lsclak_3	0.522*	0.174	0.136	How often do you ask students to make a conjecture or hypothesis from a set of given conditions?
lsclak_4	0.656*	-0.007	-0.097	How often do you ask students to solve a problem where they have to choose or hypothesize some parameters before the problem can be solved?
lsclak_5	0.147	0.574*	0.117	How often do you ask students to write proofs?
lsclak_6	0.303*	0.339*	0.182	How often do you ask students to construct mathematical arguments (e.g., justifying a solution or claim)?
lsclak_7	0.105	0.673*	-0.025	How often do you ask students to use previously established results (e.g., propositions, lemmas, theorems) to construct an argument?
lscopen_1	0.171	-0.009	0.664*	How often do you task students with problems where there are multiple solutions?
lscopen_2	0	0.002	0.907*	How often do you give students problems to solve that have more than one possible solution?
lscopen_3	0.221	0.016	0.460*	How often do you give students problems that can be solved more than one way?
lscopen_4	0.343*	-0.001	0.367*	How often do you give students a problem when you have not already told them what steps to take to solve it?
lscopen_5	0.298*	0.229*	0.285*	How often do you give students a sequence of tasks to solve that will lead them to discover something?
lsclak_8	0.875*	-0.253	0.08	How often do you ask students to critique solutions or proofs?
lsclak_9	0.284*	0.456*	0.009	How often do you give a statement to students and ask them to prove whether it is true or false?
lsclak_10	0.842*	-0.114	-0.012	How often do you provide students with arguments for them to critique?
lsclak_11	0.876*	-0.205	-0.126	How often do you ask a student to find an error in a finished proof or solution?

lsclak_12	0.284*	0.487*	0.05	How often do you ask students to use counterexamples to show that something is not true?
lsclak_13	0.533*	0.470*	-0.057	How often do you ask students to describe the idea of a proof, instead of or in addition to having them go over the details?
lsclak_14	0.500*	0.502*	-0.063	How often do you ask students to explain the big idea of a proof?

*Variables are named *xyz_#* where x=l if question pertained to a lower division course, u if question pertained to an upper division course, y=s for student, c for content, and t for teacher, z=initially hypothesized construct, and # is an arbitrary number assigned to each question pertaining to the same hypothesized construct

Example: lsclak_8 = the 8th lower division student-content item on Lakatosian practices

Table A.2: Items relevant to student-content interactions from the INQUIRE instrument and their confirmed factor structure or reason for discard

Note: Reasons for discard come from analysis during the EFA

Item	Factor or reason for discard	Statement
lscdef_1	Construct/Critique	How often do you ask students to revise a definition?
lscdef_2	Construct/Critique	How often do you ask students to propose a definition?
lscdef_3	Construct/Critique	How often do you ask students to critique a proposed definition?
lscdef_4	Prove	How often do you ask students to figure out how to use a definition?
lscdef_5	High correlation with lscdef_4 (.77)	How often do you ask students to figure out when to apply a definition?
lsclak_1	Construct/Critique	How often do you ask students to generalize a claim?
lsclak_2	Construct/Critique	How often do students propose a hypothesis for a theorem?
lsclak_3	Construct/Critique	How often do you ask students to make a conjecture or hypothesis from a set of given conditions?
lsclak_4	Construct/Critique	How often do you ask students to solve a problem where they have to choose or hypothesize some parameters before the problem can be solved?
lsclak_5	Prove	How often do you ask students to write proofs?
lsclak_6	Prove	How often do you ask students to construct mathematical arguments (e.g., justifying a solution or claim)?
lsclak_7	Prove	How often do you ask students to use previously established results (e.g., propositions, lemmas, theorems) to construct an argument?
lscopen_1	Open problems	How often do you task students with problems where there are multiple solutions?
lscopen_2	High correlation with lscopen_1 (.74)	How often do you give students problems to solve that have more than one possible solution?
lscopen_3	Open problems	How often do you give students problems that can be solved more than one way?
lscopen_4	Open problems	How often do you give students a problem when you have not already told them what steps to take to solve it?
lscopen_5	Open problems	How often do you give students a sequence of tasks to solve that will lead them to discover something?
lsclak_8	Construct/Critique	How often do you ask students to critique solutions or proofs?
lsclak_9	Prove	How often do you give a statement to students and ask them to prove whether it is true or false?
lsclak_10	Construct/Critique	How often do you provide students with arguments for them to critique?
lsclak_11	Construct/Critique	How often do you ask a student to find an error in a finished proof or solution?
lsclak_12	Prove	How often do you ask students to use counterexamples to show that something is not true?

lsclak_13	Prove	How often do you ask students to describe the idea of a proof, instead of or in addition to having them go over the details?
lsclak_14	Prove	How often do you ask students to explain the big idea of a proof?

Table A.3: Items relevant to teacher-student interactions from the INQUIRE instrument and their factor structure

Item	Original assigned factor	Assigned factor or reason for discard	Statement
ltslecture_1	interactive lecture	Interactive lecture	While teaching the whole class, how often do you pose open-ended mathematical questions to your students (i.e., questions that could not be answered with a one- or two-word answer)?
ltslecture_2	interactive lecture	interactive lecture	While teaching the whole class, how often do you make an effort to elicit questions from students (e.g., by having them fill out exit slips, use clickers, giving them time to think of questions they might have, etc.)?
ltslecture_3	interactive lecture	interactive lecture	While teaching the whole class, how often do you ask students how they would solve a problem or part of a problem?
ltslecture_4	interactive lecture	High covariance with ltslecture_1, ltslecture_7, and ltslecture_8	While teaching the whole class, how often, after demonstrating how to solve a problem, do you ask students to try a similar problem?
ltslecture_5	interactive lecture	interactive lecture	How often do you alter your lecture in response to a student's question or comment (e.g., go over an extra example, spend more time on a topic)?
ltslecture_6	interactive lecture	interactive lecture	While teaching the whole class, how often do you check to see if students are following your lesson?
ltslecture_7	interactive lecture	High covariance with ltslecture_4, ltslecture_8	While teaching the whole class, how often do you pause your presentation to ask students to work on a problem or problems?

ltslecture_8	interactive lecture	High covariance with ltslecture_3, ltslecture_4, and ltslecture_7	While you are solving a problem or constructing a proof with the whole class, how often do you ask students for suggestions of what to do next?
ltshint_1	hint	hint	If a student is stuck on a problem and asks for help during class, how often do you give them a hint on how to proceed?
ltshint_2	hint	hint	If a student is stuck on a problem and asks for help during class, how often do you try to point the student in the right direction without telling the solution (e.g., by reminding the student of a problem statement or content covered in a previous class)?
ltshint_3	hint	hint	If a student asks you to look at his or her work during class, how often do you respond without evaluating whether or not it was correct?
ltshint_4	hint	hint	If a student is stuck on a problem and asks for help during class, how often do you refer them to material already covered in the class?
ltshint_5	hint	hint	If a student is stuck on a problem and asks for help during class, how often do you help them by reminding them of an approach or strategy they've already learned?

Student-Student Interaction

Table A.4: Items relevant to student-student interactions from the INQUIRE instrument and their factor structure

Item	Original assigned factor	Assigned factor or reason for discard	Statement
lsspresent_1	Student presentations	Student presentations	How often do you have students present work to the class?
lsspresent_2	Student presentations	Student presentations	How often do you have students give feedback to student-presenters?
lsspresent_3	Student presentations	Student presentations	How often do you have students ask questions to student-presenters?
lsspresent_4	Student presentations	High cross loading on group work	How often do you have students take leadership roles during class?
lsspresent_5	Student presentations	Student presentations	How often do students come up to the board?
lsspresent_6	Student presentations	Student presentations	How often do you ask a student to study and present a new topic to the class?
lsspresent_7	Student presentations	Student presentations	How often do you ask a student to study and present a new topic to a small group of peers?
lssgroup_1	Group work	Group work	How often do you have students work together in groups?
lssgroup_2	Group work	High correlation with lssgroup_1	How often do you have students discuss a problem with each other?
lssgroup_3	Group work	Group work	If a student asks a question, how often do you redirect the question to other students?
lssgroup_4	Group work	Group work	How often do you ask students to answer another student's question?
lssgroup_5	Group work	Group work	How often do you encourage students to question each other's reasoning?
lssgroup_6	Group work	Crossloads on present and high correlation with lssgroup_1	How often do you ask students to compare their work with each other?
lssgroup_7	Group work	Group work	How often do you ask students to explain their thinking to other students?
lssgroup_8	Group work	Group work	How often do you ask students to critique mathematical arguments of their peers?
lssgroup_9	Group work	Group work	How often do you have students comment on or restate someone else's question or comment?

Teacher-Content Interaction

Table A.5: Items relevant to teacher-content interactions from the INQUIRE instrument and their factor structure

Item	Original assigned factor	1 worksheets	2 design and search for problems	3 experiences learning	Statement
ltcprep_1	Class preparation	1.396*	-0.008	0.01	How often do you prepare worksheets for students to work on during class?
ltcprep_2	Class preparation	0.063	0.661*	0.006	How often do you search for or create problems to give your students that will help them understand the course content?
ltcprep_3	Class preparation	0.024	0.642*	0.366*	How often do you design a sequence of problems so that students will discover something?
ltcprep_4	Class preparation	0.311*	0.495*	-0.093	How often do you prepare problems for students to work on during class?
ltcprep_5	Class preparation	0.342*	0.123	0.373*	How often do you design a worksheet to guide students through a difficult proof problem?
ltcprep_6	Class preparation	-0.026	0.443*	0.033	How often do you search in textbooks (including the one you're teaching from, if you are) or other resources to find material that will help students learn the course content?
ltcprep_7	Class preparation	-0.077	0.697*	0.307*	How often do you design or search for problems or activities that aim to guide students to discover something you want them to learn?
ltcprep_8	Class preparation	-0.008	-0.002	0.944*	How often do you design your lesson to include experiences you have had learning mathematics?
ltcprep_9	Class preparation	0.004	-0.017	0.964*	How often do you design your lesson to include experiences you have had doing mathematics?

* Significant at the 5% level

Appendix B: Beliefs instrument

Table A.6: Assigned factors or reason for discard for the CFA of the beliefs instrument (Clark et al., 2014), n=243

Item	Original assigned factor	Assigned factor or reason for discard	Statement
q1_1	struggle	struggle	During mathematics class, students should be asked to solve problem and complete activities by relying on their own thinking without teachers modeling an approach.
q1_2	struggle	struggle	Students can figure out how to solve many mathematics problems without being told what to do.
q1_3	struggle	struggle	During mathematics class, I do not necessarily answer students' questions immediately but rather let them struggle and puzzle things out for themselves.
q1_4	struggle	struggle	Students learn mathematics best by working to solve accessible problems that entail a solution process that has not been demonstrated to them.
q1_5	struggle	modeling	To teach mathematics, first model the activity, then provide some practice and immediate feedback, and, finally, clarify what the assignment is and how it is to be completed.
q1_6	struggle	modeling	During mathematics class, discussion should focus on students' ideas and approaches, no matter whether their answers are correct and incorrect.
q1_7	struggle	modeling	Students learn mathematics best by paying attention when their teacher demonstrates what to do, by asking question if they do not understand, and then by practicing.
q1_8	modeling	modeling	Mathematics skills are mastered incrementally, so instruction should only focus on one skill at a time, ordered by difficulty, and not move on until most students have mastered that skill.
q1_9	modeling	modeling	I like my students to master basic mathematical operations before they tackle complex problems.
q1_10	modeling	modeling	Learning mathematics requires a good memory because you must remember how to carry out procedures and, when solving an application problem, you have to remember which procedure to use.
q1_11	modeling	High correlation with q1_10 (.51)	A lot of things in mathematics must simply be accepted as true and remembered.
q1_12	modeling	modeling	When planning mathematics lessons, teachers need to focus explicitly on rules and procedures.

q1_13	modeling	modeling	Students should be homogeneously grouped for instruction and assigned to a curriculum on the basis of their prior mathematical performance.
q1_14	awareness	awareness	I learn about my students' perceptions of what 'doing mathematics' means through explicitly asking them (e.g., students write about it, one-on-one discussions, group discussions).
q1_15	awareness	awareness	I learn about my students' perceptions of connections between mathematics and their everyday lives through explicitly asking them (e.g., students write about it, one-on-one discussions, group discussions).
q1_16	awareness	awareness	I learn about my students' perceptions of connections between mathematics and their everyday lives through explicitly asking them (e.g., students write about it, one-on-one discussions, group discussions).
q1_17	awareness	High correlation with q1_18 (0.51)	For the majority of my students, I have a good sense of their motivations for wanting to succeed in mathematics.
q1_18	awareness	awareness	For the majority of my students, I have a good sense of whether or not they see how the mathematics we do in class connects to their everyday lives.
q1_19	awareness	Cross loaded on to both other factors	In order to prepare students for assessments, when students are working on a problem in mathematics, I highlight more than one approach to solving that problem.
q1_20	awareness	Cross loaded on to both other actors	I like to use mathematics problems that can be solved in many different ways.
q1_21	awareness	High correlation with q1_21 (0.42)	I have a good sense of what my unsuccessful students perceive as challenges to their mathematical performance.

Appendix C: PROSE instrument – Attending to Individual Students

This appendix begins with the introductory statement to the PROSE instrument and includes a table with the statements associated with each item and their assigned factors for a unidimensional CFA and a multidimensional EFA. Depictions of the scenarios associated with each statement are provided after the table.

Introduction
Welcome to the Attending to Individual Students survey!

This experience asks you to consider 16 teaching scenarios depicted using storyboards of cartoon characters. Each scenario shows an instructor who has decided to deviate from his or her lesson to attend to the needs of an individual student. After you view each scenario, you will be asked to indicate the extent to which you think such action is justifiable.

Please answer the questions on each screen and then click the arrow button at the bottom right of your screen.

Storyboard Instructions
In the following storyboard we invite you to consider a scenario in which a college instructor deviates from the lesson in order to attend to the needs of an individual student. We are interested in the extent to which you think the instructor's action is justifiable.

You can move through the storyboard at any rate you like. Use the arrows at the bottom left of the window to move between slides. Be sure to view the entire story.

When you finish viewing the storyboard, we will ask you to answer some questions about the scenario. You will be asked to rate the extent to which you agree with the actions taken by the instructor, and to comment on your rating. As you answer each question you will have the storyboard available to review.

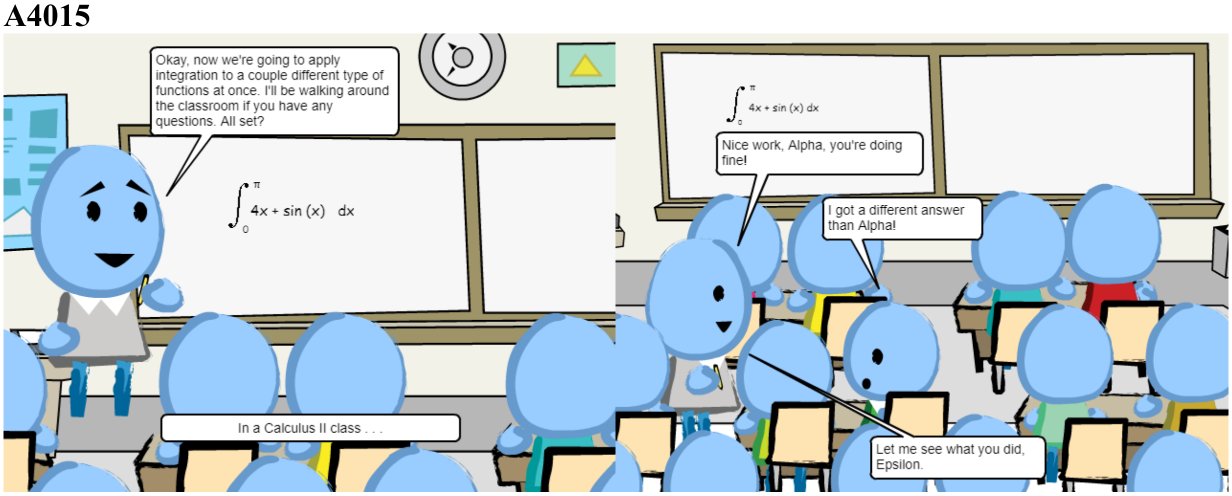
Table A.7: Item details for the unidimensional CFA for the PROSE-individual instrument

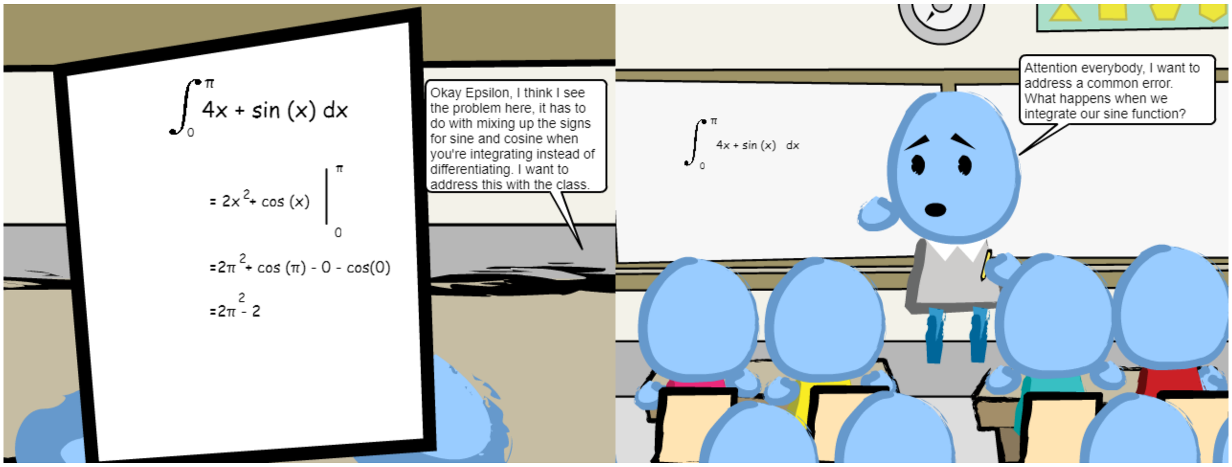
Item*	Loading or reason for discard	Statement
A4015	High correlation with A4025	The teacher should address the student’s error in private rather than discuss the error with the whole class.
A4025	Low loading	The teacher should address the student's error in private rather than discuss the error with the whole class.
A4035	Low item-rest correlation	The teacher should erase the board and go on to the next activity rather than give the student more time to copy down the problem.
A4045	Low loading	The teacher should make the student continue to work at the same pace as the rest of the class rather than let her work ahead of her peers.
A4055	Low item-rest correlation	The teacher should move on with the lesson rather than explain what the meniscus is.
A4065	Low loading	The teacher should have the student work with someone else for the class period rather than help the student with the technology.
A4075	.59	The teacher should have the student work with the same set of tools as the rest of the class rather than allow her to use an aid.
A4085	.60	The teacher should collect the quizzes in order to move on to the homework review rather than let the student have more time.
A4095	.49	The teacher should acknowledge the student and continue with the lesson rather than allow the student to come to the front and demonstrate his method for the rest of the class.
A4105	.55	The teacher should show how to solve the problem more efficiently to the whole class rather than only addressing it with the individual student.
A4115	.29	The teacher should ask the student to take the quiz with the rest of the class rather than allow the student to take the quiz another time.
A4125	low loading	The teacher should complete the example without addressing the inattentive student rather than have the student move to the front of the room and restart the example.
A4135	High correlation with A401L	The teacher should let the student manage the migraine on his own rather than give contact information for the health center.
A4145	Low item-rest correlation	The teacher should wait to speak with the student's advisor about the lack of cooperation rather than immediately lower the student's participation grade.

A4155	High correlation with A4095	The teacher should continue with the original topic rather than follow up on the aside brought up by the student.
A4165	.39	The teacher should attend to students' work on task rather than address a student's drawing ability.
A401L	.29	The instructor should remind the student of methods that were learned in class, rather than take time to learn the student's method.
A402L	.48	The instructor should help the student understand the context of the original problem, rather than change the context of the problem for that student.
A403L5	.60	The instructor should grade all quizzes in the same way, rather than offer to grade on student's quiz on the basis of what the student was able to finish.
A404L	.52	The instructor should assign the students additional work that extends what they are learning, rather than ask them what they would like to do.
A405L	.71	The instructor should ask the student to focus on the work at hand, rather than adjust lessons to engage a particular student.
A407L	.27	The instructor should insist that everyone works together in small groups, rather than allow the student to work alone.

*Naming of the items: The items are named A#%$, where A = that the item belongs to the PROSE instrument, # = the type of professional obligation, %=the number specifically associated to the item in this set, and $=the level of item, where 5 signifies college level and L signifies a linking item

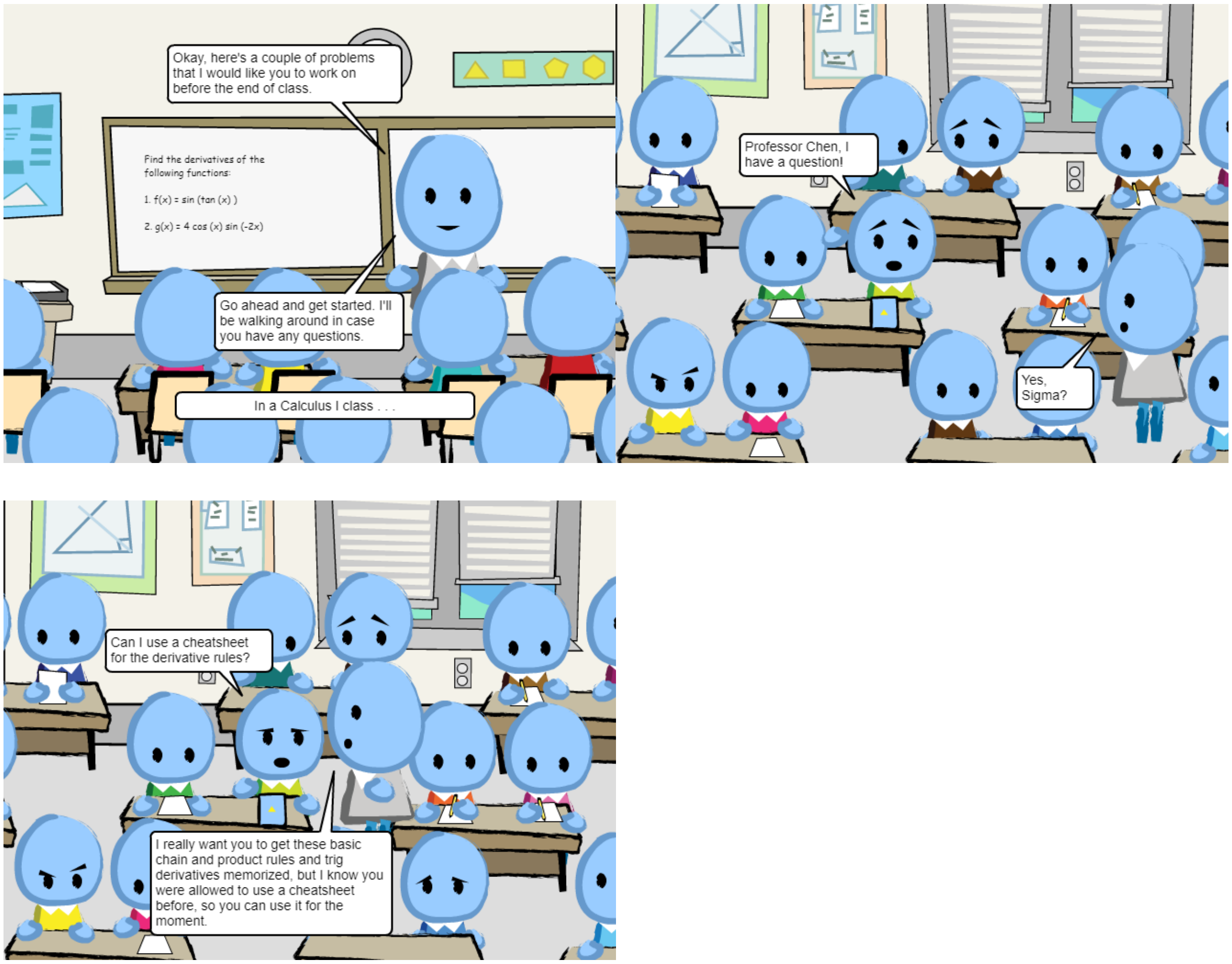
A4015





The instructor should address the student's error in private rather than discuss the error with the whole class.

A4075



The instructor should have the student work with the same set of tools as the rest of the class rather than allow her to use an aid.

Appendix D: PROSE instrument – Attending to Classroom Communities

This appendix includes a table with the statements associated with each item and their assigned factors. Depictions of the scenarios associated with each statement are provided after the table.

Table A.8: Item details for the CFA for the PROSE-interpersonal instrument

Item*	Factor loading or reason for discard	Statement
A3015	.43	The teacher should have stopped the interruptions earlier, rather than let the argument go on for so long.
A3025	.52	The teacher should encourage Beta to keep participating, rather than seek an answer from a different student in the class.
A3035	.73	The teacher should answer the student's question, rather than have him share it with the whole class.
A3045	.64	The teacher should allow the student to continue, rather than ask him to repeat his work for the rest of the class.
A3055	.65	The teacher should review the other problems on the board, rather than get more students to talk about Kappa's solution.
A3065	.67	The teacher should let the students join their small groups, rather than make them resolve their disagreement politely.
A3075	.37	The teacher should stay and answer the student's question, rather than walk away to work with others in the class.
A3085	Redundant with A306L and A3075	The teacher should let the students sit wherever they like, rather than ask them to relocate.

A3095	Low loading (.24)	The teacher should either allow or not allow Kappa to take the midterm early, rather than have her check with the rest of the class.
A3105	.61	The teacher should allow the student to work on her homework, rather than have the student help her classmate.
A3115	.44	The teacher should call on someone who has their hand in the air, rather than call on someone who has not participated yet.
A3125	Low loading (.23)	The teacher should have the activity outside as originally planned, rather than move the location to the gym.
A3135	.47	The teacher should allow the students to work with the groups they have chosen, rather than make changes after they have formed.
A3145	.34	The teacher should stay with the original example, rather than modify it.
A3155	.43	The teacher should allow the student to continue working out the problem, rather than ask her to start it over in order to make it larger.
A3165	.49	The teacher should give the mementos to the first students who asked, rather than draw sticks.
A3175	Low item-rest correlation (.18)	The teacher should let the student sharpen their pencil, rather than pause to tell them to sit down.

A3185	Low item-rest correlation (-.02)	The teacher should be lenient when students speak to each other, rather than enforce silence in the room.
A301L	Redundant with A3115 and A305L	The instructor should ask someone with the correct answer to share at the board, rather than occupy students' attention with an erroneous solution.
A302L	.67	The instructor should allow students to respond to the work on the board, rather than interrupt that discussion to ask the students to consider how they might politely phrase their comments
A303L	.60	The instructor should move on to another problem or activity, rather than insist that students share their work with a partner.
A304L	.59	The instructor should allow students that are done with their classwork to move on to the homework for the evening, rather than encourage them to check their answers with each other.
A305L	.54	The instructor should ask a student who used a traditional method to explain the solution, rather than ask the student at the board to repeat the explanation for the class to understand the alternative method.
A306L	.53	The instructor should allow the initial volunteers to share their answers, rather than continue to ask if anyone else has an answer to the question.

*Naming of the items: The items are named A#%$, where A = that the item belongs to the PROSE instrument, # = the type of professional obligation, %=the number specifically associated to the item in this set, and $=the level of item, where 5 signifies college level and L signifies a linking item

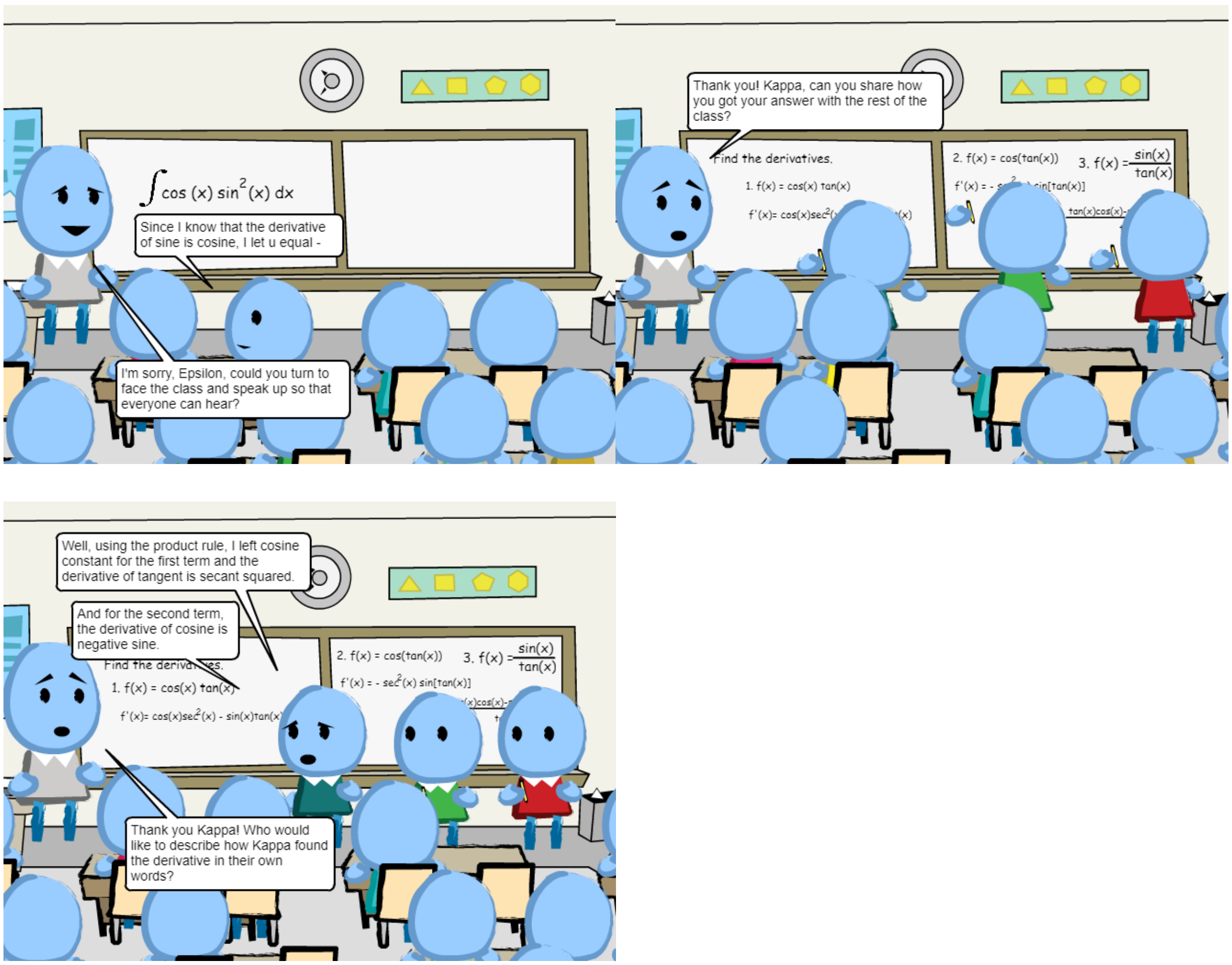
Table A.9: Item factor loadings from the EFA for the PROSE-interpersonal instrument

Item	1 Safe Environment within Classroom (not embarrassing)	2 Creating a fair/equitable experience	3 Attending to students' engagement with each other	Statement
A3015	0.252*	0.105	0.175	The teacher should have stopped the interruptions earlier, rather than let the argument go on for so long.
A3025	0.034	0.275*	0.331*	The teacher should encourage Beta to keep participating, rather than seek an answer from a different student in the class.
A3035	0.664*	0.098	0.161	The teacher should answer the student's question, rather than have him share it with the whole class.
A3045	0.258*	0.451*	0.111	The teacher should allow the student to continue, rather than ask him to repeat his work for the rest of the class.
A3055	0.544*	-0.009	0.257*	The teacher should review the other problems on the board, rather than get more students to talk about Kappa's solution.
A3065	0.12	0.251*	0.444*	The teacher should let the students join their small groups, rather than make them resolve their disagreement politely.
A3075	-0.028	0.488*	0.046	The teacher should stay and answer the student's question, rather than walk away to work with others in the class.
A3085	0.009	0.816*	-0.057	The teacher should let the students sit wherever they like, rather than ask them to relocate.
A3095	0.413*	-0.108	-0.039	The teacher should either allow or not allow Kappa to take the midterm early, rather than have her check with the rest of the class.
A3105	0.006	0.135	0.591*	The teacher should allow the student to work on her homework, rather than have the student help her classmate.
A3115	0.587*	-0.023	0	The teacher should call on someone who has their hand in the air, rather than call on someone who has not participated yet.

A3125	0.204*	-0.067	0.141	The teacher should have the activity outside as originally planned, rather than move the location to the gym.
A3135	0.098	0.175	0.307*	The teacher should allow the students to work with the groups they have chosen, rather than make changes after they have formed.
A3145	0.025	0.335*	0.087	The teacher should stay with the original example, rather than modify it.
A3155	0.195	0.352*	0.027	The teacher should allow the student to continue working out the problem, rather than ask her to start it over in order to make it larger.
A3165	0.286*	0.359*	-0.001	The teacher should give the mementos to the first students who asked, rather than draw sticks.
~~A3175~~				~~The teacher should let the student sharpen their pencil, rather than pause to tell them to sit down.~~
~~A3185~~	~~Low item-rest~~			~~The teacher should be lenient when students speak to each other, rather than enforce silence in the room.~~
A301L	0.821*	0.012	-0.194	The instructor should ask someone with the correct answer to share at the board, rather than occupy students' attention with an erroneous solution.
A302L	-0.016	0.400*	0.460*	The instructor should allow students to respond to the work on the board, rather than interrupt that discussion to ask the students to consider how they might politely phrase their comments
A303L	0.025	0.008	0.669*	The instructor should move on to another problem or activity, rather than insist that students share their work with a partner.
A304L	-0.06	0.122	0.636*	The instructor should allow students that are done with their classwork to move on to the homework for the evening, rather than encourage them to check their answers with each other.
A305L	0.476*	0.246*	-0.014	The instructor should ask a student who used a traditional method to explain the solution, rather than ask the student at the board to repeat the explanation for the class to understand the alternative method.

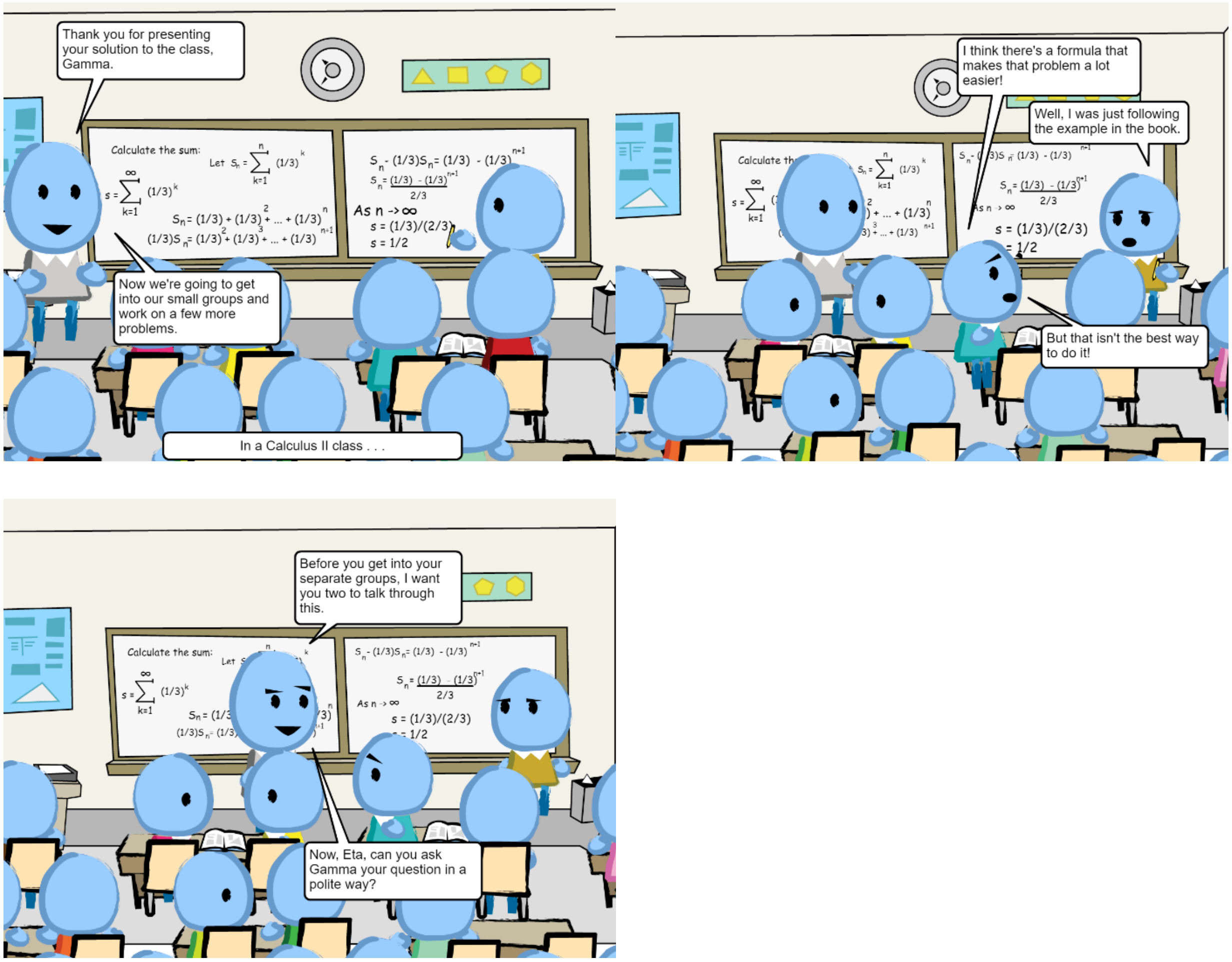
A306L	0.486*	-0.270*	0.401*	The instructor should allow the initial volunteers to share their answers, rather than continue to ask if anyone else has an answer to the question.

A3055



The instructor should review the other problems on the board, rather than get more students to talk about Kappa's solution.

A3065



The instructor should let the students join their small groups, rather than make them resolve their disagreement.

Appendix E: PROSE Instrument – Attending to Mathematical Issues

This appendix includes a table with the statements associated with each item and their assigned factors. Depictions of the scenarios associated with each statement are provided after the table.

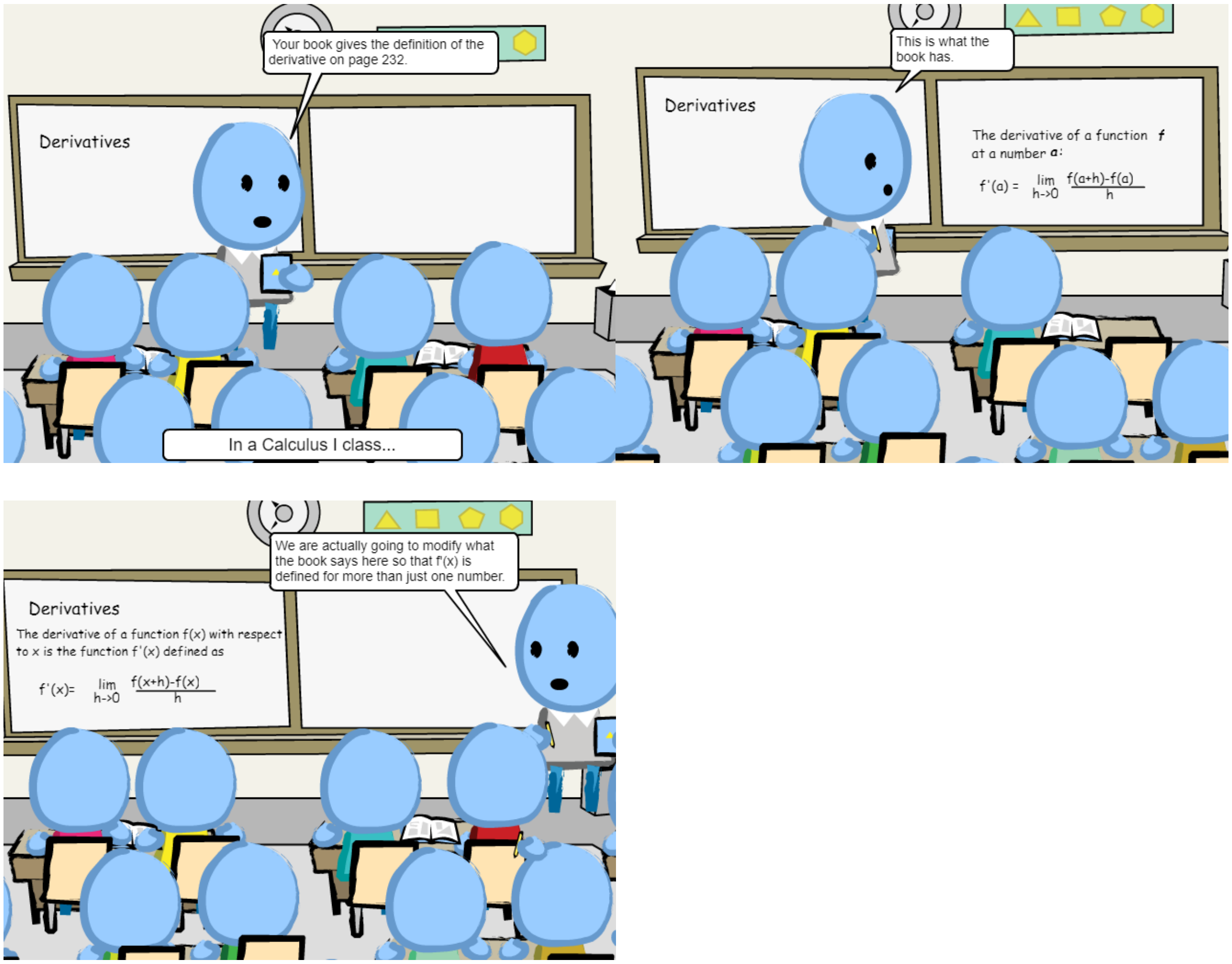
Table A.10: Item details from the unidimensional CFA for the PROSE-disciplinary instrument

Item*	Factor loadings or reason for discard	Statement
A1015	.56	The teacher should correct the mistake observed, rather than ask a student to build on mistaken work.
A1025_V2	.69	The teacher should keep to what is in the textbook, rather than require students to use information that is different from what is in the textbook.
A1035_V2	low item-rest correlation (-.03)	The teacher should explain the rationale for a convention, rather than say that the definition is arbitrary.
A1045_V2	.55	The teacher should thank the student for contributing then ask the class how to do the problem another way, rather than build on a non-standard method.
A1055	.54	The teacher should move to the next topic, rather than elaborate on details of a topic.
A1065	.56	The teacher should emphasize the method they are learning in class, rather than encourage the use of an alternative method.
A1075	low item-rest correlation (.19)	The teacher should correct the student's mathematical expression, rather than continue on with the problem.
A1085	.57	The teacher should help check the work that was finished, rather than pay attention to the idea from a student who had not yet finished.
A1095	.64	The teacher should conform to the terms that are in the textbook, rather than promote the use of a different term.
A1105	.55	The teacher should have the students work on the original problem, rather than modify the problem to better fit a real world situation.
A1115	.42	The teacher should stick to the mathematics at hand, rather than take class time to make connections to other mathematical ideas.
A1125	.66	The teacher should give students problems with specific numbers to practice, rather than ask students to produce a generalization.
A1135	.58	The teacher should give the student credit for the correct answer on the problem and move on, rather than dig further into the theory.
A1145	.69	The teacher should give students additional practice problems, rather than elaborate on mathematical theory.
A1155	High correlation with A104L (.60)	The instructor should explore the mathematics the class is actually studying, rather than bring in another mathematician to discuss her work.
A101L	.57	The teacher should assign students a problem similar to what they have been doing in class, rather than have the students continue working on the real world problem.

A102L	.32	The teacher should show a solution to the problem that uses a more traditional method, rather than ask the students to keep thinking about what makes Green's method elegant.
A103L	.38	The teacher should allow the other students who are volunteering to share their solutions at the board, rather than focus on Alpha's elegant solution.
A104L	Redundant with two other items (A1115, A102L)	The teacher should begin the day's lesson, rather than spend time considering an interesting problem.
A105L	.51	The teacher should keep the definition as originally written on the board, rather than adjust the definition in order to make it more general.
A106L	.48	The teacher should confirm that the student's method is appropriate, rather than ask the student to consider if the method would work in all cases.

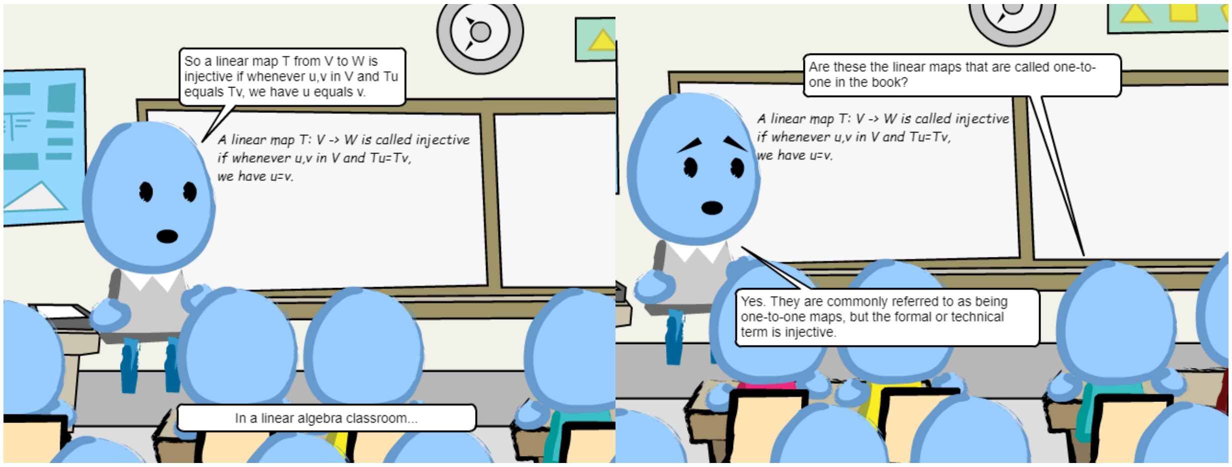
*Naming of the items: The items are named A#%$, where A = that the item belongs to the PROSE instrument, # = the type of professional obligation, %=the number specifically associated to the item in this set, and $=the level of item, where 5 signifies college level and L signifies a linking item

A1025



The instructor should keep to what is in the textbook, rather than require students to use information that is different from what is in the textbook.

A1095



The instructor should conform to the terms that are in the textbook, rather than promote the use of a different term.

Appendix F: PROSE instrument – Attending to Institutional Policy

This appendix includes a table with the statements associated with each item and their assigned factors. Depictions of the scenarios associated with each statement are provided after the table.

Table A.11: Item details for the CFA for the PROSE-Institution instrument, n=217

Item*	Reason for discard	Statement
A2015	Negative item-rest correlation (-.17)	The teacher should have students leave to accommodate the incoming class, rather than encourage them to stay longer.
A2025	.48	The teacher should answer all of the student questions, rather than stick to those in the syllabus.
A2035	Negative item-rest correlation (-.21)	The teacher should let the students remain in class, rather than let them attend the talk hosted by the university.
A2045	Negative item-rest correlation (-.08)	The teacher should ask the students to share at another time, rather than give up instructional time to discuss a school activity unrelated to mathematics.
A2055	Low loading (.172)	The teacher should allow the student to stay, rather than tell the student to enroll another semester.
A2065	Low loading (.170)	The teacher should go over the worksheet with the class, rather than use that time to go over course information.
A2075	Low loading (.24)	The teacher should let the students complete the activity without the computer, rather than insist that they use the software provided.
A2085_V2	High covariance with A201L	The teacher should let the student turn in their work, rather than require them to repeat the activity online to ensure compliance with the department's homework policy.
A2095	High covariance with A2075, A2185, A202L	The teacher should help the student by letting them view a past final, rather than restrict resources to comply with university policy.
A2105	Low loading (.24)	The teacher should honor the student's desire to remain in the class, rather than require her to attend a different mathematics class because of university policy.
A2115	.60	The teacher should answer the students' questions about related rates, rather than ignore it to keep pace with the other classes.
A2125	.43	The teacher should let the students redo the activity, rather than rely on the report from the guest professor.
A2135	Low loading (.129)	The teacher should continue with the example, rather than go back to the regular curriculum due to the presence of the department chair.

A2145	Low loading (.094)	The teacher should continue on with class material, rather than use class time to talk about a historical topic.
A2155	Low loading (.156)	The teacher should let the student remain in the room with the rest of the class, rather than insist that she leave in order to meet the requirements of the department policy.
A2165	.42	The teacher should let the student share the math video, rather than deny the request because of department rules that restrict graphing calculator usage.
A2175	.33	The teacher should wait to test students until they have spend enough time on the topic rather than give them a test in order to give the exam at the same time as the other sections.
A2185	High covariance with A2055 and A205L5	The teacher should delay the homework deadline in order to have the activity that was promised first, rather than canceling the activity.
A201L	.36	The teacher should go over the topics in the set of review materials, rather than ask the students to read them on their own.
A202L	.54	The teacher should take time to answer the students' questions, rather than move faster through the material so that the class catches up with the other sections.
A203L5	Low loading (0.234)	The teacher should allow the student at the board to continue solving the problem, rather than interrupt the work to ask the student to put down the graphing calculator.
A204L	.41	The teacher should use the remaining days to provide detailed explanations for as much of the new material as time permits, rather than briefly summarize the remaining topics.
A205L5	.50	The teacher should continue to discuss the problem as long as students have questions about it, rather than dismiss those questions because the problem will not appear on the test.
A206L5	Low item-rest correlation (0)	The teacher should help the students complete the original set of problems, rather than provide them with a different assignment.

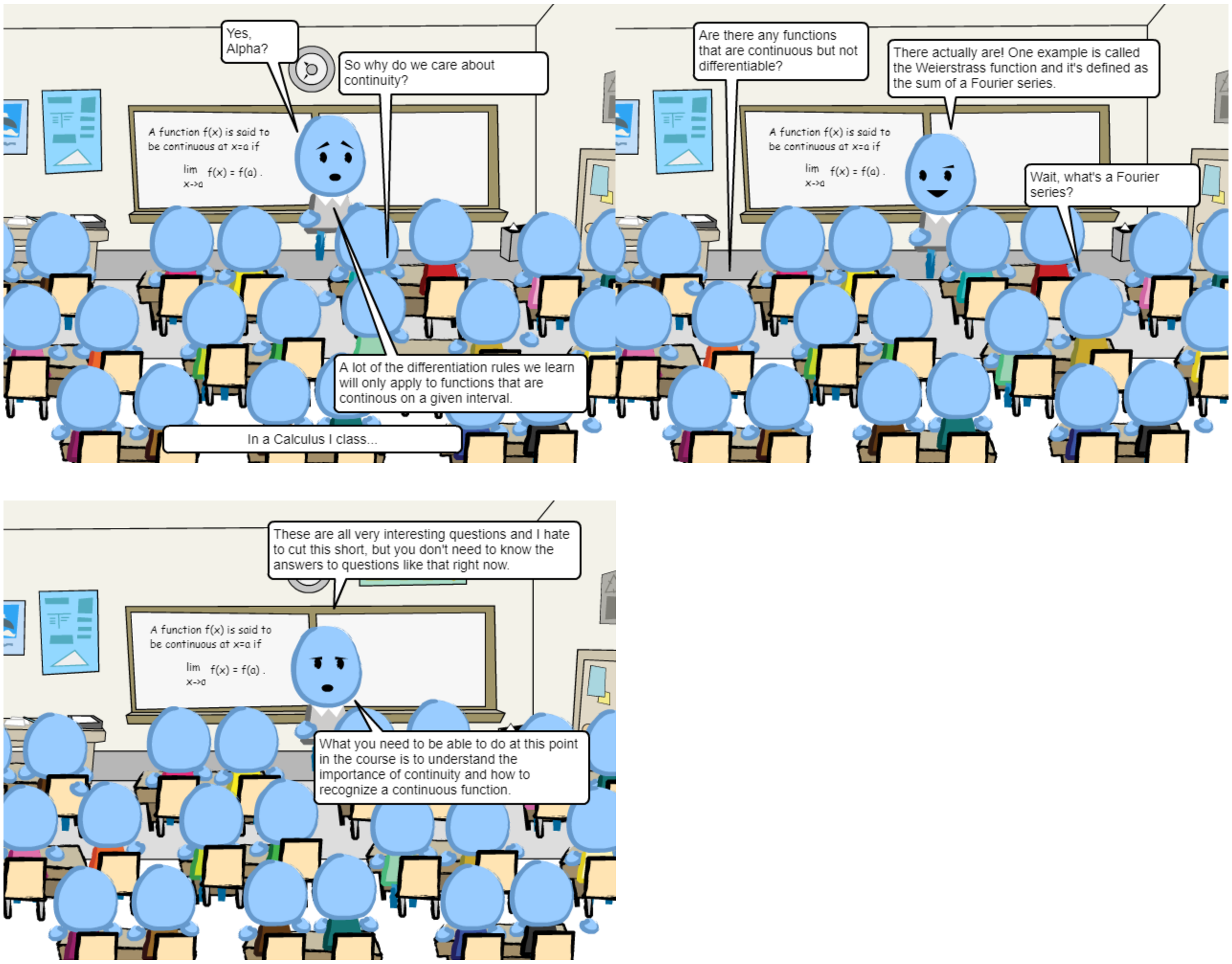
*Naming of the items: The items are named A#%$, where A = that the item belongs to the PROSE instrument, # = the type of professional obligation, %=the number specifically associated to the item in this set, and $=the level of item, where 5 signifies college level and L signifies a linking item

Table 0.12: Factor loadings for the EFA of the PROSE-Institution instrument, n=217

Item	1 Restricting	2 Pace	3	Statement
~~A2015~~	~~Low item-rest~~			~~The teacher should have students leave to accommodate the incoming class, rather than encourage them to stay longer.~~
A2025	0.165*	0.215*	0.330*	The teacher should answer all of the student questions, rather than stick to those in the syllabus.
A2035	Low item-rest			The teacher should let the students remain in class, rather than let them attend the talk hosted by the university.
A2045	Low item-rest			The teacher should ask the students to share at another time, rather than give up instructional time to discuss a school activity unrelated to mathematics.
A2055	0.101	0.226*	-0.007	The teacher should allow the student to stay, rather than tell the student to enroll another semester.
A2065	0.146*	0.1	-0.027	The teacher should go over the worksheet with the class, rather than use that time to go over course information.
A2075	0.506*	-0.317*	0.322*	The teacher should let the students complete the activity without the computer, rather than insist that they use the software provided.
A2085_V2	0.505*	-0.004	0.051	The teacher should let the student turn in their work, rather than require them to repeat the activity online to ensure compliance with the department's homework policy.
A2095	0.802*	-0.159	-0.021	The teacher should help the student by letting them view a past final, rather than restrict resources to comply with university policy.
A2105	0.058	0.056	0.261*	The teacher should honor the student's desire to remain in the class, rather than require her to attend a different mathematics class because of university policy.
A2115	-0.009	0.601*	0.178	The teacher should answer the students' questions about related rates, rather than ignore it to keep pace with the other classes.
A2125	0.306*	0.308*	0.028	The teacher should let the students redo the activity, rather than rely on the report from the guest professor.
A2135	-0.107	0.002	0.402*	The teacher should continue with the example, rather than go back to the regular curriculum due to the presence of the department chair.

A2145	0.386*	0.002	-0.423*	The teacher should continue on with class material, rather than use class time to talk about a historical topic.
A2155	-0.017	-0.209	0.594*	The teacher should let the student remain in the room with the rest of the class, rather than insist that she leave in order to meet the requirements of the department policy.
A2165	0.326*	0.022	0.400*	The teacher should let the student share the math video, rather than deny the request because of department rules that restrict graphing calculator usage.
A2175	-0.01	0.456*	-0.094	The teacher should wait to test students until they have spend enough time on the topic rather than give them a test in order to give the exam at the same time as the other sections.
A2185	0.546*	0.028	0.068	The teacher should delay the homework deadline in order to have the activity that was promised first, rather than canceling the activity.
A201L	0.444*	0.196*	-0.074	The teacher should go over the topics in the set of review materials, rather than ask the students to read them on their own.
A202L	0.033	0.416*	0.283*	The teacher should take time to answer the students' questions, rather than move faster through the material so that the class catches up with the other sections.
A203L5	0.269*	0.025	0.07	The teacher should allow the student at the board to continue solving the problem, rather than interrupt the work to ask the student to put down the graphing calculator.
A204L	0.146*	0.241*	0.108	The teacher should use the remaining days to provide detailed explanations for as much of the new material as time permits, rather than briefly summarize the remaining topics.
A205L5	-0.011	0.319*	0.289*	The teacher should continue to discuss the problem as long as students have questions about it, rather than dismiss those questions because the problem will not appear on the test.
~~A206L5~~				~~The teacher should help the students complete the original set of problems, rather than provide them with a different assignment.~~

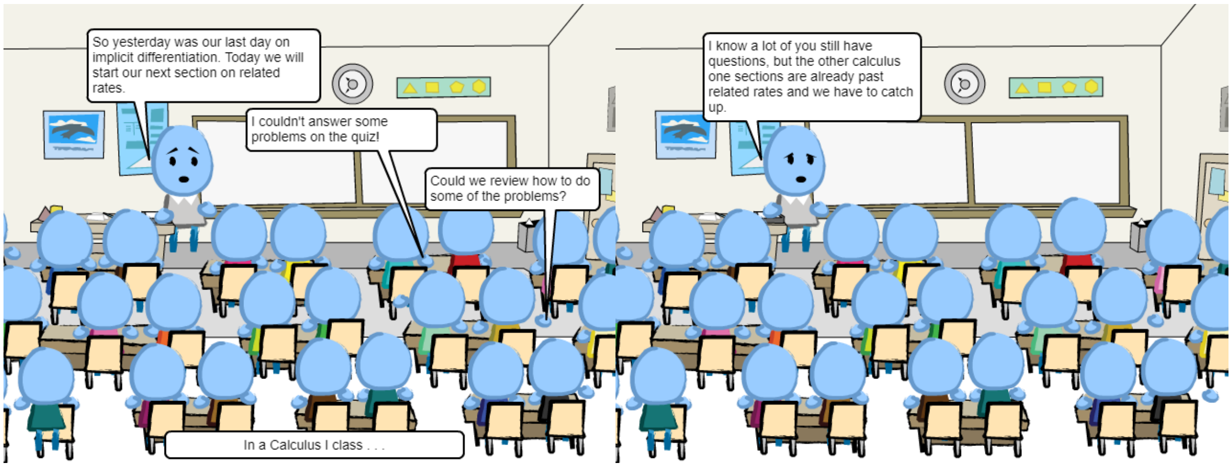
A2025



The instructor should answer all of the student questions, rather than stick to the topics in the syllabus.

The instructor should honor the student's desire to remain in the class, rather than require her to attend a different mathematics class because of university policy.

A2115



The instructor should answer the students' questions about related rates, rather than ignore them to keep pace with the other classes.

Appendix G: Background instrument

Note that many of the questions in the background instrument are conditional, that is, there are some questions participants will only see if they answer a previous question a certain way.

Closed response options in italics.

1. Please indicate your current gender identity. *Male; female; Another identity; Prefer not to answer*
2. Please indicate the category or categories with which you most identify (you may select more than one). *American Indian or Alaska Native; Asian; Black or African American; Hispanic or Latinx; Middle Eastern or North African; Native Hawaiian or Other Pacific Islander; White; Another identity; Prefer not to answer*
3. Please indicate the state where you teach.
4. What is the Carnegie Classification of your university? *R1: Highest Research Activity; R2: Higher Research Activity; R3: Moderate Research Activity; Unsure*
5. What is the highest level of formal education you have completed? *Associate's degree or equivalent; Bachelor's degree or equivalent; Master's degree; Doctoral degree; Other*
6. Do you do research? If so, specify your research interest(s).
7. By the end of the academic year, how many years will you have been teaching?
8. Which role describes your position in the department? *Graduate student; Postdoctoral fellow, Part-time faculty, Non-tenure-track faculty, Tenure-track faculty, Tenured faculty;*

None of the above

9. How many papers have you published in refereed research journals in the past two academic years (Fall 2016 – Spring 2018)?
10. What best describes the ratio of time you commit to mathematics research compared with teaching?
11. How many courses do you typically teach per semester (quarter/trimester)?
12. What courses are you currently teaching?
13. What is a typical enrollment for one section for the smallest course you typically teach?
14. Have you heard of inquiry-oriented or inquiry-based learning?
    a. [If **yes** to 14] What does inquiry-oriented instruction mean to you?
    b. [If **yes** to 14] Have you taken a course taught with inquiry-oriented instruction (including but not limited to inquiry-based learning or the Modified Moore Method)?
    c. [If yes to 14c] What type of course(s) did you take with inquiry-oriented instruction?
    d. [If **yes** to 14] Anything else you'd like to share about your experience taking an inquiry-oriented course or courses? (E.g., who you took it with or what the course(s) were like.)
    e. [If **yes** to 14] Have you ever taught with inquiry-oriented or inquiry-based learning?
    f. [If **yes** to 14e] In what courses have you used inquiry-oriented or inquiry-based learning?
    g. [If **yes** to 14g] Are you encouraged to use inquiry-oriented or inquiry-based learning by your department or institution?
15. If you would like to be considered for an Amazon gift card drawing once you have completed all 7 questionnaires (including this one), please fill in the following information.

We will not use this information or any other part of the study. (*Email, address, city, state, postal code, and country.)*

Appendix H: Descriptive Statistics

The maximum possible value of some questions (all variables that begin with lts- or u-) was 5 due to an error in the initial survey design. The missingness shown by decrease in observations is for three reasons: (1) general survey fatigue, (2) not everyone taught an upper-division course, and (3) if participants selected the option 1-Never in response to the question "How often do you ask students to revise a definition?" they were skipped past all remaining question blocks (all variables that begin with lss-, ltc- and u-). The third reason was due to an error in survey implementation where a "skip" was unintentionally inserted in the survey logic. The participants that selected 1-Never in response to a question about how often they asked students to craft definitions were skipped past the remainder of the survey. These participants were asked to complete the survey again after it was fixed, approximately half complied.

Table A.13: Descriptive statistics for items in the INQUIRE instrument, with respect to lower-division courses

Variable	*n*	Mean	Std. Dev.	Min	Max
ltslecture_1	257	3.11	1.18	1	5
ltslecture_2	257	3.42	1.3	1	5
ltslecture_3	257	4.02	1.05	1	5
ltslecture_4	257	3.59	1.17	1	5
ltslecture_5	257	3.19	1.03	1	5
ltslecture_6	257	4.31	0.91	1	5
ltslecture_7	257	3.61	1.30	1	5
ltslecture_8	257	4.05	1.09	1	5
ltshint_1	257	3.83	0.97	1	5
ltshint_2	257	3.92	1.06	1	5
ltshint_3	257	2.49	1.26	1	5
ltshint_4	257	3.16	1.06	1	5
ltshint_5	257	3.83	0.89	1	5
lscdef_1	252	1.7	1.02	1	6
lscdef_2	252	1.9	1.13	1	5
lscdef_3	252	1.79	1.10	1	6
lscdef_4	252	2.88	1.42	1	6
lscdef_5	252	3.02	1.39	1	6
lsclak_1	252	2.31	1.33	1	6
lsclak_2	252	1.80	1.04	1	5
lsclak_3	252	2.38	1.31	1	6
lsclak_4	252	2.15	1.21	1	6
lsclak_5	252	1.82	1.16	1	6
lsclak_6	252	3.28	1.53	1	6
lsclak_7	252	3.01	1.53	1	6
lscopen_1	252	2.83	1.29	1	6
lscopen_2	252	2.89	1.24	1	6
lscopen_3	252	3.92	1.21	1	6
lscopen_4	252	3.09	1.39	1	6
lscopen_5	252	3.03	1.33	1	6
lsclak_8	252	2.18	1.25	1	6
lsclak_9	252	2.65	1.30	1	6
lsclak_10	252	1.97	1.15	1	6
lsclak_11	252	2.17	1.16	1	6
lsclak_12	252	2.82	1.25	1	6
lsclak_13	252	2.09	1.29	1	6
lsclak_14	252	2.11	1.28	1	6
lsspresent_1	194	2.17	1.32	1	6

lsspresent_2	194	1.68	1.11	1	6
lsspresent_3	194	1.88	1.29	1	6
lsspresent_4	194	2.14	1.29	1	6
lsspresent_5	194	2.29	1.39	1	6
lsspresent_6	194	1.38	0.89	1	5
lsspresent_7	194	1.49	1.03	1	6
lssgroup_1	194	3.85	1.73	1	6
lssgroup_2	194	4.10	1.64	1	6
lssgroup_3	194	2.93	1.44	1	6
lssgroup_4	194	3.03	1.50	1	6
lssgroup_5	194	2.90	1.64	1	6
lssgroup_6	194	3.67	1.78	1	6
lssgroup_7	194	3.54	1.58	1	6
lssgroup_8	194	2.24	1.42	1	6
lssgroup_9	194	2.38	1.43	1	6
ltcprep_1	194	3.62	1.85	1	6
ltcprep_2	194	4.75	1.39	1	6
ltcprep_3	194	3.28	1.49	1	6
ltcprep_4	194	4.62	1.57	1	6
ltcprep_5	194	2.45	1.52	1	6
ltcprep_6	194	4.35	1.58	1	6
ltcprep_7	194	3.66	1.62	1	6
ltcprep_8	194	3.30	1.67	1	6
ltcprep_9	194	3.13	1.62	1	6

Table A.14: Descriptive statistics for items in the INQUIRE instrument, with respect to upper-division courses

Variable	*n*	Mean	Std. Dev.	Min	Max
utslecture_1	79	3.82	1.17	1	5
utslecture_2	79	3.80	1.26	1	5
utslecture_3	79	3.96	1.13	1	5
utslecture_4	79	3.41	1.20	1	5
utslecture_5	79	3.35	1.04	1	5
utslecture_6	79	4.29	1.09	1	5
utslecture_7	79	3.58	1.26	1	5
utslecture_8	79	4.15	1.10	1	5
utshint_1	79	3.78	1.11	1	5
utshint_2	79	3.73	1.14	1	5
utshint_3	79	2.51	1.21	1	5
utshint_4	79	3.43	1.12	1	5
utshint_5	79	3.63	1.10	1	5
uscdef_1	79	2.11	1.13	1	5
uscdef_2	79	2.23	1.13	1	5
uscdef_3	79	2.32	1.23	1	5
uscdef_4	79	3.35	1.14	1	5
uscdef_5	79	3.33	1.25	1	5
usclak_1	79	2.82	1.21	1	5
usclak_2	79	2.33	1.15	1	5
usclak_3	79	2.77	1.17	1	5
usclak_4	79	2.11	1.14	1	5
usclak_5	79	3.32	1.41	1	5
usclak_6	79	3.72	1.27	1	5
usclak_7	79	3.66	1.22	1	5
uscopen_1	79	3.27	1.14	1	5
uscopen_2	79	3.14	1.14	1	5
uscopen_3	79	3.56	1.07	1	5
uscopen_4	79	3.13	1.20	1	5
uscopen_5	79	3.05	1.06	1	5
usclak_8	79	2.54	1.22	1	5
usclak_9	79	3.10	1.01	1	5
usclak_10	79	2.48	1.11	1	5
usclak_11	79	2.28	1.07	1	5
usclak_12	79	3.16	1.18	1	5
usclak_13	79	2.86	1.25	1	5
usclak_14	79	2.86	1.25	1	5
usspresent_1	79	2.11	1.24	1	5

usspresent_2	79	1.75	1.07	1	5
usspresent_3	79	1.89	1.13	1	4
usspresent_4	79	1.92	1.22	1	5
usspresent_5	79	2.35	1.27	1	5
usspresent_6	79	1.53	0.97	1	5
usspresent_7	79	1.63	1.05	1	5
ussgroup_1	79	3.06	1.50	1	5
ussgroup_2	79	3.30	1.44	1	5
ussgroup_3	79	2.76	1.23	1	5
ussgroup_4	79	2.71	1.22	1	5
ussgroup_5	79	2.75	1.31	1	5
ussgroup_6	79	3.08	1.58	1	5
ussgroup_7	79	3.16	1.41	1	5
ussgroup_8	79	2.30	1.25	1	5
ussgroup_9	79	2.37	1.17	1	5
utcprep_1	79	2.68	1.59	1	5
utcprep_2	79	4.03	1.09	1	5
utcprep_3	79	3.10	1.29	1	5
utcprep_4	79	3.56	1.44	1	5
utcprep_5	79	2.34	1.34	1	5
utcprep_6	79	4.06	1.16	1	5
utcprep_7	79	3.52	1.31	1	5
utcprep_8	79	3.14	1.26	1	5
utcprep_9	79	3.08	1.32	1	5

Table A.15: Descriptive statistics for items in the beliefs instrument, n=243

Variable	Mean	Std. Dev.	Min	Max
1	3.85	1.17	1	6
2	3.89	1.31	1	6
3	4.05	1.24	1	6
4	4.095	1.11	1	6
5	3.65	1.16	1	6
6	3.84	1.33	1	6
7	3.79	1.33	1	6
8	3.18	1.27	1	6
9	4.47	1.13	2	6
10	2.81	1.30	1	6
11	2.16	1.19	1	6
12	2.60	1.23	1	5
13	3.05	1.32	1	6
14	3.33	1.42	1	6
15	3.38	1.41	1	6
16	3.55	1.40	1	6
17	4.16	1.06	1	6
18	3.60	1.13	1	6
19	4.54	1.02	1	6
20	4.65	1.02	1	6
21	4.04	1.06	1	6

Table A.16: Descriptive statistics for items in the PROSE-interpersonal instrument

Variable	Obs	Mean	Std. Dev.	Min	Max
A3095	205	2.26	1.26	1	6
A3015	205	3.98	1.42	1	6
A3115	205	3.58	1.22	1	6
A304L	205	4.14	1.16	1	6
A3145	205	4.44	1.27	1	6
A306L	205	3.08	1.32	1	6
A3075	205	3.92	1.23	1	6
A3125	205	3.79	1.27	1	6
A3055	205	3.66	1.38	1	6
A301L	205	3.41	1.39	1	6
A3105	205	4.12	1.35	1	6
A305L	205	4.29	1.05	1	6
A3085	205	4.44	1.18	1	6
A3185	205	2.87	1.30	1	6
A3035	205	3.87	1.36	1	6
A3025	205	3.99	1.20	1	6
A3065	205	3.81	1.34	1	6
A3045	205	3.79	1.49	1	6
A3165	205	4.18	1.34	1	6
A3175	205	3.23	1.34	1	6
A302L	205	4.15	1.38	1	6
A3135	205	3.11	1.32	1	6
A303L	205	4.24	1.33	1	6
A3155	205	3.76	1.34	1	6

Table A.17: Descriptive statistics for the items from the PROSE-disciplinary instrument, n=209

Variable	Mean	Std. Dev.	Min	Max
A1115	4.55	1.27	1	6
A1045	4.13	1.28	1	6
A104L	4.12	1.3	1	6
A1155	4.19	1.3	1	6
A1125	4.94	1.13	1	6
A1025	4.75	1.12	1	6
A1145	4.62	1.14	1	6
A1085	4.74	1.18	1	6
A1015	3.88	1.35	1	6
A1035	2.02	0.97	1	5
A1065	4.38	1.24	1	6
A102L	3.06	1.19	1	6
A105L	3.78	1.19	1	6
A1075	2.67	1.33	1	6
A1105	4.55	1.35	1	6
A101L	4.89	1.13	1	6
A106L	4.22	1.26	1	6
A103L	3.1	1.25	1	6
A1095	4.32	1.22	1	6
A1055	4.34	1.21	1	6
A1135	4.02	1.19	1	6

Table A.18: Descriptive statistics for the items from the PROSE-institution instrument, n=217

Variable	Mean	Std. Dev.	Min	Max
A203L5	3.38	1.39	1	6
A2035	4.15	1.3	1	6
A2075	3.98	1.46	1	6
A2095	5.25	1.13	1	6
A2185	4.15	1.38	1	6
A2135	2.45	1.28	1	6
A2055	3.12	1.56	1	6
A202L	2.61	1.2	1	6
A2015	2.93	1.43	1	6
A2155	2.9	1.42	1	6
A201L	4.08	1.28	1	6
A205L5	2.73	1.26	1	6
A2065	3.28	1.46	1	6
A2175	3.12	1.47	1	6
A2115	3.07	1.28	1	6
A2165	3.56	1.53	1	6
A204L	3.3	1.25	1	6
A2085	4.18	1.49	1	6
A206L5	2.81	1.43	1	6
A2025	3.63	1.36	1	6
A2145	4.69	1.2	1	6
A2105	2.81	1.36	1	6
A2125	3.85	1.24	1	6
A2045	2.98	1.51	1	6

Table A.19: Descriptive statistics for the items from the PROSE-individual instrument

Variable	Obs	Mean	Std. Dev.	Min	Max
A4165	236	2.88	1.46	1	6
A4115	236	3.77	1.41	1	6
A4125	236	2.62	1.54	1	6
A4015	236	4.6	1.19	1	6
A407L	236	4	1.37	1	6
A4155	236	3.31	1.54	1	6
A4145	236	3.21	1.66	1	6
A4105	236	2.7	1.36	1	6
A402L	235	3.36	1.4	1	6
A4085	235	2.38	1.2	1	6
A404L	235	3.32	1.45	1	6
A4075	234	2.88	1.54	1	6
A4095	234	2.79	1.65	1	6
A4045	234	4.62	1.32	1	6
A4035	234	3.97	1.43	1	6
A401L	234	4.96	1.16	1	6
A405L	234	3	1.44	1	6
A4055	234	5.12	1.15	1	6
A4065	234	3.99	1.32	1	6
A403L5	234	1.75	1.2	1	6

For the sake of creating mean descriptive scores in Table A.20, I rescaled items that had a 5-point scale to a 6-point scale.

Table A.20: Mean values for constructs from all instruments

Instrument	Construct	*n*	Mean	Std. Dev.	Min	Max
INQUIRE (lower-division)	Active lecture	275	4.39	0.84	2	6
	Hinting without telling	275	4.29	0.97	1	6
	Constructing/Critiquing	270	2.06	0.86	1	4.9
	Proving	270	2.6	0.96	1	5.38
	Open problems	270	3.24	0.94	1	6
	Group work	212	3.01	1.22	1	6
	Student presentations	212	1.83	0.92	1	5.5
Beliefs	Struggle	243	3.94	0.87	1.6	6
	Model	243	3.36	0.8	1.43	5.43
	Awareness	243	3.46	1.02	1	5.75
PROSE	Institution	217	3.33	0.68	1.44	5.33
	Discipline	211	4.24	0.69	1.5	5.82
	Individual	236	3.15	0.71	1.5	5.17
	Interpersonal	205	3.76	0.62	1.75	5.25

Appendix I: Internal Consistency

The reliability statistics showing internal consistency item-by-item are reported here for the constructs that were originally hypothesized. These statistics helped guide some of the initial decision-making whether to keep items and, if so, where to group them while conducting the factor analyses.

Table A.21: Reliability statistics for the student-content definition-formulating items in the INQUIRE instrument, n=252

Item	Sign	Item-test correlation	Item-rest correlation	Average interitem correlation	alpha
lscdef_1	+	0.74	0.57	0.46	0.78
lscdef_2	+	0.78	0.63	0.44	0.76
lscdef_3	+	0.79	0.65	0.43	0.75
lscdef_4	+	0.72	0.55	0.47	0.78
lscdef_5	+	0.73	0.56	0.47	0.78
Test scale				0.45	0.81

Table A.22: Reliability statistics for the student-content constructing items in the INQUIRE instrument, n=252

Item	Sign	Item-test correlation	Item-rest correlation	Average interitem correlation	alpha
lsclak_1	+	0.79	0.69	0.41	0.8
lsclak_2	+	0.72	0.60	0.43	0.82
lsclak_3	+	0.8	0.71	0.40	0.8
lsclak_4	+	0.61	0.46	0.47	0.84
lsclak_5	+	0.67	0.53	0.45	0.83
lsclak_6	+	0.71	0.59	0.43	0.82
lsclak_7	+	0.72	0.59	0.43	0.82
Test scale				0.43	0.84

Table A..23: Reliability statistics for the student-content open problem items in the INQUIRE instrument, n=252

Item	Sign	Item-test correlation	Item-rest correlation	Average interitem correlation	alpha
lsclak_8	+	0.74	0.63	0.46	0.83
lsclak_9	+	0.69	0.57	0.47	0.84
lsclak_10	+	0.74	0.64	0.45	0.83
lsclak_11	+	0.74	0.63	0.46	0.83
lsclak_12	+	0.69	0.56	0.47	0.84
lsclak_13	+	0.76	0.66	0.45	0.83
Test scale				0.46	0.86

Table A..24: Reliability statistics for the student-content critiquing items in the INQUIRE instrument, n=252

Item	Sign	Item-test correlation	Item-rest correlation	Average interitem correlation	alpha
lsclak_1	+	0.79	0.69	0.41	0.80
lsclak_2	+	0.72	0.60	0.43	0.82
lsclak_3	+	0.8	0.71	0.40	0.80
lsclak_4	+	0.61	0.46	0.47	0.84
lsclak_5	+	0.67	0.53	0.45	0.83
lsclak_6	+	0.71	0.59	0.43	0.82
lsclak_7	+	0.72	0.59	0.43	0.82
Test scale				0.43	0.84

Table A.25: Reliability statistics for the teacher-student interactive lecture items in the INQUIRE instrument, n=252

Item	Sign	Item-test correlation	Item-rest correlation	Average interitem correlation	alpha
ltslecture_1	+	0.60	0.45	0.29	0.74
ltslecture_2	+	0.60	0.45	0.29	0.74
ltslecture_3	+	0.68	0.55	0.27	0.73
ltslecture_4	+	0.56	0.40	0.30	0.75
ltslecture_5	+	0.55	0.39	0.30	0.75
ltslecture_6	+	0.58	0.42	0.30	0.75
ltslecture_7	+	0.66	0.52	0.28	0.73
ltslecture_8	+	0.67	0.54	0.28	0.73
Test scale				0.29	0.77

Table A.26: Reliability statistics for the student-content hinting without telling items in the INQUIRE instrument, n=252

Item	Sign	Item-test correlation	Item-rest correlation	Average interitem correlation	alpha
ltshint_1	+	0.73	0.54	0.32	0.66
ltshint_2	+	0.79	0.63	0.29	0.62
ltshint_3	+	0.56	0.31	0.42	0.75
ltshint_4	+	0.63	0.40	0.38	0.71
ltshint_5	+	0.75	0.57	0.31	0.64
Test scale				0.35	0.73

Table A.27: Reliability statistics for the student-student presentation items in the INQUIRE instrument, n=194

Item	Sign	Item-test correlation	Item-rest correlation	Average interitem correlation	alpha
lsspresent_1	+	0.76	0.66	0.51	0.86
lsspresent_2	+	0.82	0.75	0.49	0.85
lsspresent_3	+	0.86	0.79	0.48	0.85
lsspresent_4	+	0.70	0.58	0.54	0.87
lsspresent_5	+	0.79	0.70	0.50	0.86
lsspresent_6	+	0.71	0.59	0.53	0.87
lsspresent_7	+	0.71	0.59	0.53	0.87
Test scale				0.51	0.88

Table A.28: Reliability statistics for the student-student group work items in the INQUIRE instrument, n=194

Item	Sign	Item-test correlation	Item-rest correlation	Average interitem correlation	alpha
lssgroup_1	+	0.77	0.70	0.58	0.92
lssgroup_2	+	0.78	0.72	0.57	0.91
lssgroup_3	+	0.80	0.74	0.57	0.91
lssgroup_4	+	0.81	0.75	0.57	0.91
lssgroup_5	+	0.84	0.79	0.56	0.91
lssgroup_6	+	0.81	0.75	0.57	0.91
lssgroup_7	+	0.82	0.77	0.56	0.91
lssgroup_8	+	0.72	0.64	0.59	0.92
lssgroup_9	+	0.73	0.65	0.59	0.92
Test scale				0.57	0.92

Table A.29: Reliability statistics for the teacher-content items in the INQUIRE instrument, n=194

Item	Sign	Item-test correlation	Item-rest correlation	Average interitem correlation	alpha
ltcprep_1	+	0.57	0.43	0.32	0.79
ltcprep_2	+	0.60	0.47	0.31	0.78
ltcprep_3	+	0.75	0.66	0.28	0.76
ltcprep_4	+	0.55	0.41	0.32	0.79
ltcprep_5	+	0.59	0.45	0.31	0.79
ltcprep_6	+	0.49	0.34	0.33	0.80
ltcprep_7	+	0.69	0.58	0.29	0.77
ltcprep_8	+	0.66	0.54	0.30	0.77
ltcprep_9	+	0.67	0.55	0.30	0.77
Test scale				0.31	0.80

Table A.30: Reliability statistics for the TASSP items from the beliefs instrument, n=243

Item	Sign	Item-test correlation	Item-rest correlation	Average interitem correlation	alpha
q1_1	+	0.75	0.57	0.35	0.69
q1_2	+	0.75	0.58	0.35	0.68
q1_3	+	0.67	0.47	0.40	0.72
q1_4	+	0.68	0.48	0.39	0.72
q1_6	+	0.69	0.49	0.39	0.72
Test				0.38	0.75

Table A.31: Reliability statistics for the TMIM items from the beliefs instrument, n=243

Item	Sign	Item-test correlation	Item-rest correlation	Average interitem correlation	alpha
q1_7	+	0.73	0.57	0.32	0.70
q1_8	+	0.65	0.46	0.35	0.73
q1_9	+	0.66	0.48	0.34	0.72
q1_10	+	0.71	0.54	0.33	0.71
q1_11	+	0.59	0.39	0.37	0.75
q1_12	+	0.69	0.52	0.33	0.71
Test scale				0.34	0.76

Table A.32: Reliability statistics for the TASMD items from the beliefs instrument, n=243

Item	Sign	Item-test correlation	Item-rest correlation	Average interitem correlation	alpha
q1_14	+	0.69	0.56	0.31	0.75
q1_15	+	0.70	0.58	0.30	0.75
q1_16	+	0.69	0.56	0.30	0.75
q1_17	+	0.61	0.47	0.32	0.77
q1_18	+	0.58	0.43	0.33	0.78
q1_19	+	0.61	0.46	0.32	0.77
q1_20	+	0.61	0.46	0.32	0.77
q1_21	+	0.58	0.43	0.33	0.78
Test scale		0	0	0.32	0.79

Table A.33: Reliability statistics for items from the PROSE-individual instrument

Item	*n*	Sign	Item-test correlation	Item-rest correlation	Average interitem correlation	alpha
A4165	236	+	0.48	0.38	0.12	0.75
A4115	236	+	0.36	0.26	0.13	0.76
A4125	236	+	0.31	0.21	0.13	0.76
A4015	236	+	0.41	0.31	0.13	0.75
A407L	236	+	0.37	0.27	0.13	0.76
A4155	236	+	0.48	0.39	0.12	0.75
A4145	236	+	0.27	0.16	0.13	0.76
A4105	236	+	0.38	0.28	0.13	0.75
A402L	235	+	0.44	0.34	0.13	0.75
A4085	235	+	0.47	0.37	0.12	0.75
A404L	235	+	0.43	0.33	0.13	0.75
A4075	234	+	0.48	0.39	0.12	0.75
A4095	234	+	0.51	0.42	0.12	0.75
A4035	234	+	0.34	0.23	0.13	0.76
A401L	234	+	0.46	0.37	0.12	0.75
A405L	234	+	0.57	0.49	0.12	0.74
A4055	234	+	0.34	0.24	0.13	0.76
A4045	234	+	0.21	0.1	0.14	0.77
A4065	234	+	0.37	0.27	0.13	0.76
A403L5	234	+	0.42	0.32	0.13	0.75
A4025	234	+	0.39	0.28	0.13	0.75
A4135	233	+	0.48	0.39	0.12	0.75
Test scale					0.13	0.76

Table A.34: Reliability statistics for items from the PROSE-interpersonal instrument, n=205

Item	Sign	Item-test correlation	Item-rest correlation	Average interitem correlation	alpha
A3095	+	0.27	0.19	0.20	0.85
A3015	+	0.43	0.36	0.20	0.85
A3115	+	0.48	0.41	0.19	0.85
A304L	+	0.58	0.52	0.19	0.84
A3145	+	0.38	0.30	0.20	0.85
A306L	+	0.55	0.48	0.19	0.84
A3075	+	0.42	0.35	0.20	0.85
A3125	+	0.28	0.20	0.20	0.85
A3055	+	0.62	0.56	0.19	0.84
A301L	+	0.51	0.44	0.19	0.84
A3105	+	0.58	0.51	0.19	0.84
A305L	+	0.52	0.46	0.19	0.84
A3085	+	0.54	0.47	0.19	0.84
A3185	+	-0.02	-0.11	0.22	0.86
A3035	+	0.70	0.65	0.18	0.84
A3025	+	0.53	0.46	0.19	0.84
A3065	+	0.60	0.54	0.19	0.84
A3045	+	0.61	0.55	0.19	0.84
A3165	+	0.53	0.46	0.19	0.84
A3175	+	0.18	0.10	0.21	0.86
A302L	+	0.60	0.54	0.19	0.84
A3135	+	0.49	0.42	0.19	0.85
A303L	+	0.56	0.50	0.19	0.84
A3155	+	0.47	0.40	0.19	0.85
Test scale				0.19	0.85

Table A.35: Reliability statistics for items from the PROSE-disciplinary instrument

Item	n	Sign	Item-test correlation	Item-rest correlation	Average interitem correlation	alpha
A1115	211	+	0.48	0.40	0.22	0.85
A1045	211	+	0.54	0.46	0.21	0.84
A104L	210	+	0.58	0.51	0.21	0.84
A1155	210	+	0.55	0.47	0.21	0.84
A1125	210	+	0.62	0.55	0.21	0.84
A1025	209	+	0.62	0.55	0.21	0.84
A1145	209	+	0.62	0.55	0.21	0.84
A1085	209	+	0.53	0.46	0.21	0.84
A1015	209	+	0.54	0.47	0.21	0.84
A1035	209	+	0.06	-0.03	0.24	0.86
A1065	209	+	0.55	0.48	0.21	0.84
A102L	209	+	0.42	0.34	0.22	0.85
A105L	209	+	0.52	0.44	0.21	0.84
A1075	209	+	0.28	0.19	0.23	0.85
A1105	209	+	0.55	0.48	0.21	0.84
A101L	209	+	0.54	0.46	0.21	0.84
A106L	209	+	0.49	0.42	0.21	0.85
A103L	209	+	0.46	0.38	0.22	0.85
A1095	209	+	0.57	0.5	0.21	0.84
A1055	209	+	0.50	0.42	0.21	0.85
A1135	209	+	0.56	0.49	0.21	0.84
Test scale					0.21	0.85

Table A.36: Reliability statistics for items from the PROSE-institution instrument, n=217

Item	Sign	Item-test correlation	Item-rest correlation	Average interitem correlation	alpha
A2075	+	0.35	0.22	0.05	0.55
A2095	+	0.42	0.29	0.05	0.54
A2185	+	0.45	0.32	0.05	0.53
A2055	+	0.26	0.12	0.06	0.56
A202L	+	0.44	0.32	0.05	0.54
A2035	+	-0.07	-0.21	0.07	0.61
A2015	+	-0.03	-0.17	0.06	0.6
A203L5	+	0.32	0.19	0.05	0.55
A2135	+	0.23	0.1	0.06	0.57
A201L	+	0.49	0.38	0.05	0.53
A205L5	+	0.39	0.26	0.05	0.54
A2115	+	0.48	0.36	0.05	0.53
A2165	+	0.42	0.29	0.05	0.54
A2175	+	0.25	0.11	0.06	0.56
A2145	+	0.24	0.10	0.06	0.57
A206L5	+	0.14	0.00	0.06	0.58
A2045	+	0.06	-0.08	0.06	0.59
A2155	+	0.2	0.06	0.06	0.57
A204L	+	0.38	0.25	0.05	0.55
A2085	+	0.44	0.32	0.05	0.54
A2025	+	0.42	0.29	0.05	0.54
A2105	+	0.37	0.23	0.05	0.55
A2125	+	0.43	0.30	0.05	0.54
Test scale				0.05	0.57

References

Andersen, M. (2011). *Knowledge, attitudes, and instructional practices of Michigan community college math instructors: The search for a knowledge, attitudes, and practices gap in collegiate mathematics,* (Doctoral dissertation). Western Michigan University, Kalamazoo, MI.

Austin, A. E. (2002). Preparing the next generation of faculty: Graduate school as socialization to the academic career. *The Journal of Higher Education, 73*(1), 94-122.

Ball, D. L. (1993). With an eye on the mathematical horizon: Dilemmas of teaching elementary school mathematics. *The Elementary School Journal*, 93(4): 373-397.

Ball, D. L., & Forzani, F. M. (2007). 2007 Wallace Foundation Distinguished Lecture - What makes Education Research "Educational"? *Educational Researcher, 36*(9), 529-540.

Banchi, H. & Bell, R. (2008). The many level of inquiry: Inquiry comes in various forms. *Science and Children, 46*(2), 26.

Barkatsas, A. & Malone, J. (2005). A typology of mathematics teachers' beliefs about teaching and learning mathematics and instructional practices. *Mathematics Education Research Journal, 17*(2), 69-90.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

Bentler, P. M., & Chou, C. (1987). Practical issues in structural equation modeling. *Sociological Methods and Research, 16,* 78-117.

Blair, R., Kirkman, E. E., & Maxwell, J. W. (2018). Statistical abstract of undergraduate programs in the mathematical sciences in the United States: Fall 2015 CBMS survey: American Mathematical Society.

Boaler, J. (1997). Reclaiming school mathematics: The girls fight back. *Gender and Education, 9*(3), 285-305.

Boaler, J. (1998). Open and closed mathematics: Student experiences and understanding. *Journal for Research in Mathematics Education, 29*(1), 41-62.

Boileau, N., Ko, I., & Herbst, P. (2018, April). Uncovering the relationship between expertise in teaching mathematics and teachers' recognition of their professional obligations. Paper presented at the Annual Meeting of AERA, New York, NY.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, NY: Guilford Press.

Bruner, J. (1960). *The process of education.* Cambridge MA: Harvard University Press.

Brousseau, G. (1997). Theory of didactical situations in mathematics: Didactique des mathématiques 1970-1990. (N. Balacheff, M. Cooper, & V. Warfield, Trans. & Eds.). Dordrecht, The Netherlands: Kluwer Academic.

Brousseau, G. & Warfield, G. (2003). Glossary of terms used in Didactique. Retrieved from http://faculty.washington.edu/warfield/guy-brousseau.com/biographie/glossaires/.

Burn, H. E., Mesa, V., & White, N. (2015). Calculus I in community colleges: Findings from the National CSPCC Study. *MathAMATYC Educator*, 6(3), 34-39.

Burton, L. (2004). *Mathematicians as enquirers: Learning about learning mathematics.* Norwell, MA: Kluwer Academic Publishers.

Cai, J., Morris, A., Hohensee, C., Hwang, S., Robison, V., Cirillo, M., Kramer, S. L., & Hiebert, J. (2019). Theoretical framing as justifying. *Journal of Research in Mathematics Education, 50*(3), 218-224.

Cai, J. & Wang, T. (2010). Conceptions of effective mathematics teaching within a cultural context: perspectives of teachers from China and the United States. *Journal of Mathematics Teacher Education, 13*(3), 265-287.

Cattell, R. B. (1978). *The scientific use of factor analysis in behavioral and life sciences.* New York, NY: Plenum Press.

Chalice, D. R. (1995). How to teach a class by the Modified Moore Method. *The American Mathematical Monthly, 102*(4), 317-321.

Chazan, D. and Herbst, P. (2011). Challenges of Particularity and Generality in Depicting and Discussing Teaching. *For the Learning of Mathematics*, *31*(1), 9-13.

Chazan, D., Herbst, P., & Clark, L. (2016). Research on the teaching of mathematics: A call to theorize the role of society and schooling in mathematics instruction. In D. Gitomer & C. Bell (Eds.), *Handbook of research on teaching* (Fifth ed., pp. 1039-1097): American Educational Research Association.

Clark, L., DePiper, J., Frank, T., Nishio, M., Campbell, P., Smith, T., Griffin, M., Rust, A., Conant, D. & Choi, Y. (2014). Teacher characteristics associated with mathematics teachers' beliefs and awareness of their students' mathematical dispositions. *Journal for Research in Mathematics Education, 45*(2), 246-284.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*(3), 309-319.

Cobb, P., Wood, T., Yackel, E., Nicholls, J., Wheatley, G., Trigatti, B., & Perlwitz, M. (1991). Assessment of problem-centered second-grade mathematics project. *Journal for Research in Mathematics Education, 22*(1), 3-29.

Cohen, D. W. (1982). A modified Moore method for teaching undergraduate mathematics. *The American Mathematical Monthly, 89*(7), 473-474+487-490.

Cohen, D. K. (1990). A revolution in one classroom: The case of Mrs. Oublier. *Education Evaluation and Policy Analysis, 12*(3), 311-329.

Cohen, D., Raudenbush, S., & Ball, D. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis, 25*(2), 119-142.

Cooney, T. J. (1985). A beginning teacher's view of problem solving. *Journal for Research in Mathematics Education, 16*, 324–336.

Cross, D. I. (2009). Alignment, cohesion and change: Examining mathematics teachers' belief structures and their influence on instructional practices. *Journal of Mathematics Teacher Education, 12*(5), 325-346.

Cross Francis, D. I. (2015). Dispelling the notion of inconsistencies in teachers' mathematics beliefs and practices: A 3-year case study. *Journal of Mathematics Teacher Education, 18*(2), 173–201. DOI: 10.1007/s10857-014-9276-5

Dane, E., & Pratt, M. G. (2007). Exploring intuition and its role in managerial decision making. *The Academy of Management Review, 32*(1), 33-54.

Davies, W. M. (2009). Groupwork as a form of assessment: common problems and recommended solutions. *Higher Education, 58*(4), 563-584.

DeFranco, T. C., & McGivney-Burelle, J. (2001). The beliefs and instructional practices of mathematics teaching assistants participating in a mathematics pedagogy course. In R.

Speiser, C. A. Maher, & C. N. Walter (Eds), *Proceedings of the Annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*, Snowbird, Utah (pp. 681-90). Columbus, OH: ERIC/CSMEE.

Desimone, L. M., Smith, T. M., & Frisvold, D. E. (2010). Survey measure of classroom instruction: Comparing student and teacher reports. *Educational Policy, 24*(2), 267-329.

DeVellis, R. F. (2016). *Scale development: Theory and applications* (Vol. 26). Newbury Park, CA: Sage publications.

Dewey, J. (1902). *The child and the curriculum.* Chicago, IL: University of Chicago Press.

Dewey, J. (1933). *How we think.* Boston: D. C. Heath.

De Vleeschouwer, M. & Gueudet, G. (2011). Secondary-tertiary transition and evolution of didactic contract: The example of duality in linear algebra. In M. Pytlak, T. Rowland, & E. Swoboda (Eds.), *Proceedings of the seventh congress of the European mathematical Society for Research in mathematics education* (pp. 2013-2022). Rzesow: University of Rzeszow and ERME.

D'Onofrio, L. (1989). The use of self-reports on sensitive behaviours in health program evaluation. *New directions in program evaluation, 49,* 59-74.

Earnest, P. (1989). The impact of beliefs on the teaching of mathematics. In P. Earnest (Ed.), *Mathematics teaching: The stateof the art* (pp. 249-254). London: Falmer Press.

Eddy, S. L., & Kogan, K. A. (2014). Getting under the hood: How and for whom does increasing course structure work? *CBE - Life Sciences Education, 13,* 453-468.

Educational Advancement Foundation. (2017). *Educational Advancement Foundation.* Retrieved from http://www.eduadvance.org .

Edwards, B. S. and Ward, M. B. (2004). Surprises from mathematics education research: Student (mis)use of mathematical definitions. *The American Mathematical Monthly, 111,* 411-424.

Engeln, K., Euler, M., & Katja, M. (2013). Inquiry-based learning in mathematics and science: a comparative baseline study of teachers' beliefs and practices across 12 European countries. *ZDM – The International Journal of Mathematics Education, 45*(6), 823-836.

Ernest, P. (1991). *The philosophy of mathematics education.* London: The Falmer Press.

Erickson, A. & Herbst, P. (in press, online first Sept 14, 2016). Will teachers create opportunities for discussion when teaching proof in a geometry classroom? *International Journal of Mathematics and Science Education*.

Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis* (5th ed.). West Sussex, United Kingdom: John Wiley & Sons.

Finch, J. (1987). The vignette technique in survey research. *Sociology, 21*(1), 105-114.

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. *Proceedings of the National Academy of Sciences of the United States of America, 111*(23): 8410-8415.

Freire, P. (1990). *Pedagogy of the oppressed.* New York: Continuum. (Original work published in 1968).

Furr, R. M. & Bacharach, V. R. (2013). *Psychometrics: An Introduction* (2nd Ed.). Thousand Oaks, CA: Sage publications.

Gellert, U. (2000). Mathematics instruction in safe space: Prospective elementary teachers' views of mathematics education. *Journal of Mathematics Teacher Education, 3*(3), 251-270.

Gholson, M. & Martin, D. B. (2014). Smart girls, black girls, mean girls, and bullies: At the intersection of identities and the mediating role of young girls' social network in mathematical communities of practice. *Journal of Education, 194*(1), 19-33.

Golbeck, A., Barr, T. H., & Rose, C. A. (2018). Fall 2016 departmental profile report: Annual survey of the mathematical sciences in the US. *Notices of the AMS*, *65*(8), 949-959.

Gonzalez, J. J. (2013). My journey with inquiry-based learning. *Journal on Excellence in College Teaching, 24*(2), 33-50.

Goodstein, D. & Neugebauer, G. (1995). Special preface. In R. Reynman (with R. B. Leighton and M. Sands), *Six easy pieces* (pp. xix-xxii), New York, NY: Basic Books.

Gutiérrez, R. (2013). The sociopolitical turn in mathematics education. *Journal for Research in Mathematics Education, 44*(1), 37-68.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). Multivariate data analysis (Vol. 6). Upper Saddle River, NJ: Pearson.

Halmos, P. (1980). The heart of mathematics. *American Mathematical Monthly, 87*, 519-524.

Hawkins, D. (1974). I, though, and it. In D. Hawkins, *The informed vision: Essays on learning and human nature* (pp. 48-62). New York: Agathon. (Original work published 1967).

Hayden, H. E., Moore-Russo, D., & Marino, M. R. (2012). One teacher's reflective journey and the evolution of a lesson: Systematic reflection as a catalyst for adaptive expertise. *International and Multidisciplinary Perspectives, 14*(1), 144-156.

Hayward, C. N., Kogan, M., & Laursen, S. L. (2016). Facilitating instructor adoption of Inquiry-Based Learning in college mathematics. *International Journal of Research in Undergraduate Mathematics Education, 2*(1), 59-82. doi:10.1007/s40753-015-0021-y

Hayward, C. N., Weston, T., & Laursen, S. L. (2018). First results from a validation study of TAMI: Toolkit for Assessing Mathematics Instruction. In A. Weinberg, C. Rasmussen, J. Rabin, M. Wawro, & S. Brown (Eds.), *Proceedings of the 21st Annual Conference on Research in Undergraduate Mathematics Education* (pp. 727-735). San Diego, CA, 2018.

Henderson, K. (1963). Research on teaching secondary school mathematics. In N. L. Gage (Ed.), *Handbook of research on teaching.* Chicago: Rand McNally.

Henriksson, C., & Friesen, N. (2012). Introduction. In N. Friesen, C. Henriksson, & T. Saevi (Eds.), *Hermeneutic phenomenology in education: Method and practice* (pp. 1-16). Sense Publishers: Boston.

Herbst, P. & Chazan, D. (2011). Research on practical rationality: Studying the justification of actions in mathematics teaching. *The Mathematics Enthusiast, 8*(3), 405–462.

Herbst, P. & Chazan, D. (2012). On the instructional triangle and sources of justification for actions in mathematics teaching. *ZDM – The International Journal of Mathematics Education, 44*(5), 601-612.

Herbst, P. & Chazan, D. (2015). Studying professional knowledge use in practice using multimedia scenarios delivered online. *International Journal of Research & Method in Education, 38*(3), 272-287.

Herbst, P., Chazan, D., Chen, C. L., Chieu, V. M., & Weiss, M. (2011). Using comic-based representations of teaching, and technology, to bring practice to teacher education courses. *ZDM – The International Journal on Mathematics Education, 49*, 91-103.

Herbst, P., Chazan, D., Kosko, K., Dimmel, J. & Erickson, A. (2016). Using multimedia questionnaires to study influences on the decisions mathematics teachers make in instructional situations. *ZDM-The international journal of mathematics education, 48,* 167-183.

Herbst, P. & Ko, I. (2018, April). Recognition of professional obligations of mathematics teaching and their role in justifying instructional actions. Paper presented at the Annual Meeting of AERA, New York, NY.

Herbst, P. & Miyakawa, T. (2008). When, how and why prove theorems? A methodology for studying the perspective of geometry teachers. *ZDM Mathematics Education, 40*, 469-486.

Herzig, A. (2002). Where have all the students gone? Participation of doctoral students in authentic mathematical activity as a necessary condition for persistence toward the Ph.D. *Educational Studies in Mathematics, 50*(2), 177-212.

Herzig, A. H. (2004). Becoming mathematicians: Women and students of color choosing and leaving doctoral mathematics. *Review of Educational Research, 74,* 1015-1019.

Hiebert, J., Carpenter, T. P., Fennema, B., Fuson, K. C., Wearne, D. & Murray, H. (1997). *Making sense: Teaching and learning mathematics with understanding*. Portmouth, NH: Heinemann.

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist, 42*(2), 99-107.

Holstein, J. A. & Gubrium, J. F. (1995). The active interview. Thousand Oaks, CA: Sage.

Holton, D. (Ed.) (2001). *The teaching and learning of mathematics at the university level.* Dordrecht: Kluwer.

Hoyles, C. (1992), Mathematics teaching and mathematics teachers: A meta-case study. *for the Learning of Mathematics*, *12*(3), 32-44.

Hu, L. T., & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal, 6*(1), 1-55.

International Association for the Evaluation of Educational Achievement. (2014). TIMSS: Teacher questionnaire advanced mathematics. Retrieved from https://nces.ed.gov/timss/pdf/2015_12th_grade_Teacher_Questionnaire_Advanced_Math.pdf

Jacobson, E. & Izák, A. (2015). Knowledge and motivation as mediators in mathematics teaching practice: the case of drawn models for fraction arithmetic. *Journal of Mathematics Teacher Education, 18*, 467-488.

Johnson, E., Andrews-Larson, C., Keene, K., Melhuish, K., Keller, R., & Fortune, N. (2018). *Inquiry and inequity in the undergraduate mathematics classroom.* Paper presented at the 40th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, Greenville, SC.

Johnson, E., Caughman, J., Fredericks, J., & Gibson, L. (2013). Implementing inquiry-oriented curriculum: From the mathematicians' perspective. *The Journal of Mathematical Behavior, 32*(4), 743-760.

Jones, F. B. (1977). The Moore Method. *American Mathematical Monthly*, 273-278.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20,* 141-151.

Karabenick, S. A., Wooley, M. E., Friedel, J. M., Ammon, B. V., Blazevski, J., Bonney, C. R., …Kelly, K. L. (2007). Cognitive processing of self-report items in educational research: Do they think what we mean? *Educational Psychologist, 42*(3), 139-151.

Kennedy, M. M. (1999). Approximations to indicators of student outcomes. Educational Evaluation and Policy Analysis, 21(4), 345–363. doi:10.3102/01623737021004345

Kerry, J. (1995). Harvard calculus at Oklahoma State University. *The American Mathematical Monthly, 102*(9). 794-797.

Kersting, N. B., Givvin, K. B., Sotelo, F. L., & Stigler, J. W. (2010). Teachers' analyses of classroom video predict student learning of mathematics: Further explorations of a novel measure of teacher knowledge. *Journal of Teacher Education, 61*(1-2), 172-181.

Kilpatrick, J., Swafford, J., & Findell, B. (Eds.) (2001). *Adding it up: Helping children learn mathematics.* Washington DC: National Academy Press.

Kim, H. S. (2002). We talk, therefore we think? A cultural analysis of the effect of talking on thinking. *Journal of Personality and Social Psychology, 83*(4), 828-842.

Kim, H. S. (2008). Culture and the cognitive and neuroendocrine responses to speech. *Journal of Personality and Social Psychology*, *94*(1), 32-47.

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th Ed.). New York, NY: The Guilford Press.

Kogan, M. & Laursen, S. L. (2014). Assessing long-term effects of inquiry-based learning: A case study from college mathematics. *Innovative Higher Education*, *39*(3), 183-199.

Kuhs, T. M., & Ball, D. L. (1986). *Approaches to teaching mathematics: Mapping the domains of knowledge, skills and disposition.* East Lansing, MI: Center on Teacher Education, Michigan State University.

Kuster, G., Johnson, E., Keene, K., & Andrews-Larson, C. (2017). Inquiry-oriented instruction: A conceptualization of the instructional principles. *Problems, Resources, and Issues in Mathematics Undergraduate Studies, 28*(1), 13-30.

Kwon, O. N., Rasmussen, C., & Allen, K. (2005). Students' retention of mathematical knowledge and skills in differential equations. *School Science and Mathematics, 105*(5), 1-13.

Lai, E. R., & Waltman, K. (2008). Test preparation: Examining teacher perceptions and practices. *Educational Measurement: Issues and Practice, 27*(2), 28-45.

Lakatos, I. (1976). *Proofs and refutations: The logic of mathematical discovery.* New York: NY: Cambridge University Press.

Lampert, M. (1985). How do teachers manage to teach? Perspectives on problems in practice. *Harvard Educational Review, 55*(2), 178-194.

Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American Education Research Journal, 27*(1), 29-63.

Lande, E., & Mesa, V. (2016). Instructional decision making and agency of community college mathematics faculty. *ZDM – The International Journal on Mathematics Education*, *48*(1-2), 199-212.

Larsen, S., Johnson, E., & Bartlo, J. (2013). Designing and scaling up an innovation in abstract algebra. *The Journal of Mathematical Behavior, 32*(4), 693-711.

Laursen, S., & Hassi, M.-L. (2010). *Benefits of inquiry based learning for undergraduate college mathematics students.* Paper presented at the Annual meeting of the American Educational Research Association, Denver, CO.

Laursen, S., Hassi, M. L., Kogan, M., & Weston, T. (2014). Benefits for women and men of inquiry-based learning in college mathematics: A multi-institution study. *Journal for Research in Mathematics Education, 45*(4), 406-418.

Laursen, S. & Rasmussen, C. (2019). I on the Prize: Inquiry approaches in undergraduate mathematics. *International Journal of Research in Undergraduate Mathematics Education, 5*(1), 129-146.

Leatham, K. (2006). Viewing mathematics teachers' beliefs as sensible systems. *Journal of Mathematics Teacher Education,9*(1), 91–102.

Leron, U. & Dubinsky, E. (1995). An abstract algebra story. *American Mathematical Monthly, 102*(3), 227-242.

Lesh, R., & Harel, G. (2003). Problem solving, modeling, and local conceptual development. *Mathematical Thinking and Learning, 5*(2&3), 157-190.

Lesh, R., & Yoon, C. (2004). Evolving communities of mind – in which development involves several interacting and simultaneously developing strands. *Mathematical Thinking and Learning, 6*(2), 205-226.

Lew, K., Fukawa-Connelly, T. P., Mejía-Ramos, J. P., & Weber, K. (2016). Lectures in advanced mathematics: Why students might not understand what the mathematics professor is trying to convey. *Journal for Research in Mathematics Education, 47*(2), 162-198.

Lewis, J. M., Fischman, D., & Riggs, M. (2015). Defining, developing, and measuring "Proclivities for Teaching Mathematics." *Journal of Mathematics Teacher Education*, *18*(5), 447–465.

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling, 9*(2), 151-173.

Lock, P. F. (1994). Reflectinos on the Harvard calculus approach. *Problems, Resources, and Issues in Mathematics Undergraduate Studies, 4*(3), 229-234.

Lockwood, E., Johnson, E., & Larsen, S. (2013). Developing instructor support materials for an inquiry-oriented curriculum. *The Journal of Mathematical Behavio*r, *32*(4), 776-790.

Lubienski, S. T. (2000). Problem solving as a means toward mathematics for all: An exploratory look through a class lens. *Journal for Research in Mathematics Education, 31*(4), 454-482.

Mahavier, W. S. (1999). What is the Moore Method? *Problems, Resources, and Issues in Mathematics Undergraduate Studies*, *9*(4), 339-354. doi:10.1080/10511979908965940a

Marshall, J. C. & Horton, R. M. (2011). The relationship of teacher-facilitated, inquiry-based intruction to student higher-order thinking. *School Science and Mathematics, 111*(3), 93-101.

Mason, J. (2001). Mathematical teaching practices at the tertiary level: Working group report. In D. A. Holton (Ed.), *The Teaching and Learning of Mathematics at University Level: An ICMI Study* (pp. 71-86). Hingham, MA: Kluwer.

Matsunaga, M. (2008). Item parceling in structural equation modeling: A primer. *Communication methods and measures, 2*(4), 260-293.

Matyas, M. L., & Dix, L. S. (1992). Science and engineering programs: On target for women? Washington, DC: National Academy Press.

McDuffie, A. R., & Graeber, A. O. (2003). Institutional norms and policies that influence colege mathematics professors in the process of changing to reform-based practices. *School Science and Mathematics, 103*(7), 331-344.

Mcgivney-Burelle, J., DeFranco, T. C., Vinsonhaler, C. I., & Santucci, K. B. (2001). Building Bridges: Improving the Teaching Practices of TAs in the Mathematics Department. *Journal of Graduate Teaching Assistant Development*, *8*(2), 55-63.

Mesa, V., Celis, S. & Lande, E. (2014). Teaching approaches of community college mathematics faculty: Do they relate to classroom practices? *American Educational Research Journal, 51(1),* 117-151.

Mesa, V., Shultz, M., & Jackson, A. (2019). Moving away from lecture in undergraduate mathematics: Managing tensions within a coordinated inquiry-based linear algebra course. International Journal for Research in Undergraduate Mathematics, Advanced online publication. DOI: /10.1007/s40753-019-00109-1

Mojena, R. (1977). Hierarchical grouping methods and stopping rules: An evaluation. *The Computer Journal, 20*(4), 359-363.

Moore-Russo, D.& Vigietti, J. M. (2010). Teachers' reactions to animations as representations of geometry instruction. *ZDM – The International Journal on Mathematics Education, 43*(1), 161-173.

Moore-Russo, D. & Wilsey, J. N. (2014). Delving into the meaning of productive reflection: A study of future teachers' reflections on representations of teaching. *Teaching and Teacher Education, 37,* 76-90.

National Center for Education Statistics. (1997). Digest of education statistics, 1997 (NCES Publication No. 98-015). Washington, DC: Author.

National Research Council. (1996). *National science education standards.* Washington, D.C.: National Academy Press.

National Research Council. (2000). *Inquiry and the national science education standards: A guide for teaching and learning*. Washington, D.C.: National Academy Press.

Neff, J. A. (1979). Interaction versus hypothetical others: The use of vignettes in attitude research. *Sociology and social research, 64*(1), 105-125.

Newfield, J. (1980). Acurracy of teacher self-reports: Reports and observations of specific classroom behaviors. *Journal of Educational Research, 74*, 78-82.

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research, 62*(3), 307-332.

Parker, J. (2005). *R. L. Moore: Mathematician and teacher.* Mathematical Association of America.

Pedaste, M., Mäeots, M., Siiman, L. A., De Jong, T., Van Riesen, S. A., Kamp, E. T., Tsourlidaki, E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational research review, 14*, 47-61.

Philipp, R. A. (2007). Mathematics teachers' beliefs and affect. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 257-315). Charlotte, NC: Information Age Publishing.

Pintrich, P. R. (1990). Implications of psychological research on student learning and college teaching or teacher education. In W. R. Houston (Ed.), *Handbook of research on teacher education* (pp. 826-857). New York: Macmillan.

Pólya, G. (1962). *Mathematical discovery: On understanding, learning and teaching problem solving.* New York: Wiley.

President's Council of Advisors on Science and Technology. (2012). Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics. Retrieved from https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/pcast-engage-to-excel-final_2-25-12.pdf.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology, 25,* 111-163.

Rasmussen, C., Apkarian, N., Hagman, J. E., Johnson, E., Larsen, S., Bressoud, D., The Progress through Calculus Team. (2019). Characteristics of precalculus through calculus 2 programs: Insights from a national census survey. *Journal for Research in Mathematics Education, 50*(1), 98-111.

Rasmussen, C., & Ellis, J. (2015). Calculus coordination at PhD-granting universities: More than just using the same syllabus, textbook, and final exam. In D. Bressoud, V. Mesa, & C. Rasmussen (Eds.), *Making the connection: Research and teaching in undergraduate mathematics education* (pp. 107-115). Washington, DC: The Mathematical Association of America.

Rasmussen, C. & Kwon, O. N. (2007). An inquiry-oriented approach to undergraduate mathematics. *Journal of Mathematical Behavior, 26*, 189-194.

Rasmussen, C., Kwon, O. N., Allen, K., Marrongelle, K., & Burtch, M. (2006). Capitalizing on advances in mathematics and K-12 mathematic education in undergraduate mathematics: An inquiry-oriented approach to differential equations. *Asia Pacific Education Review, 7*(1), 85-93.

Raykov, T. (2004). Behavioral scale reliability and measurement invariance evaluation using latent variable modeling. *Behavior Therapy, 35,* 299-331.

Raymond, A. (1997). Inconsistencies between a beginning elementary teacher's mathematics beliefs and teaching practice. *Journal for Research in Mathematics Education, 28*(5), 550–576.

Renz, P. (1999). The Moore Method: What discovery learning is and how it works. *FOCUS: Newsletter of the Mathematical Association of America, 19,* 6-8.

Rickards, G., Magee, C., & Artino, A. R. Jr. (2014). You can't fix by analysis what you've spoiled by design: developing survey instruments and collecting validity evidence. *Journal of Graduate Medical Education*, *4*(4), 407-410.

Ross, J. A., McDougall, D., Hogaboam-Gray, A., & Sage, A. (2003) A survey measuring elementary teachers' implementation of standards-based mathematics teaching.

Rowan, B., Harrison, D. M., & Hayes, A. (2004). Using instructional logs to study mathematics curriculum and teaching in the early grades. *The Elementary School Journal, 105*(1), 103-127.

Satyam, V. R. (2018). *Cognitive and affective components of undergraduate students learning how to prove (Unpublished doctoral dissertation)*. Michigan State University: East Lansing, MI.

Scheffler, I. (1965). *Conditions of knowledge: An introduction to epistemology and education.* Glenview, IL: Scott, Foresman and Company.

Schoenfeld, A. H. (1988). When good teaching leads to bad results: The disasters of "well-taught" mathematics courses. *Educational Psychologist, 23*(2), 145-166.

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

Schwab, J. J. (1961/1974). Education and the structure of the disciplines. In I. Westbury & N. Wilkof (Eds.), *Science, curriculum, and liberal education* (pp. 229-272). Chicago: University of Chicago Press.

Schwab, J. J. (1978). The practical: Translation into curriculum. In I. Westbury & N. J. Wilkof (Eds.), *Science, curriculum, and liberal education: Selected essays* (pp. 365-383). Chicago: University of Chicago Press.

Shultz, M., Herbst, P., & Schleppegrell, M. (2019). A study of agency: How professors and graduate teaching assistants position themselves in relation to their instructional obligations. *Linguistics and Education, 52*, 33-43.

Silver, E. A. (1994). On mathematical problem posing. *for the learning of mathematics, 14*(1), 19-28.

Silver, E. A. (1997). Fostering creativity through instruction rich in mathematical problem solving and problem posing. *ZDM – The International Journal of Mathematics Education*, *29*(3), 75-80.

Skott, J. (2001). The emerging practices of a novice teacher: The roles of his school mathematics images. *Journal of Mathematics Teacher Education, 4*, 3–28.

Smith, J. C. (2005). A sense-making approach to proof: Strategies of students in traditional and problem-based number theory courses. *Journal of Mathematical Behavior, 25,* 73-90.

Smith, T. M., Desimone, L. M., & Ueno, K. (2005). “Highly qualified” to do what? The relationship between NCLB teacher quality mandates and the use of reform-oriented instruction in middle school mathematics. *Educational Evaluation and Policy Analysis, 27*(1), 75–109. doi:10.3102/01623737027001075

Speer, N. (2001). *Connecting beliefs and teaching practices: A study of teaching assistants in reform-oriented calculus courses (Unpublished doctoral dissertation)*. University of California, Berkeley.

Speer, N. (2005). Issues of methods and theory in the study of mathematics teachers’ professed and attributed beliefs. *Educational Studies in Mathematics, 58*(3), 361-391..

Stains, M., Harshman, J., Barker, M. K., Chasteen, S. V., Cole, R., DeChenne-Peters, S. E., Eagan Jr., M. K., Esson, J. M., Knight, J. K., Laski, F. A., Levis-Fitzgerald, M., Lee, C. J., Lo, S. M., McDonnell, L. M., McKay, T. A., Michelotti, N., Musgrove, A., Palmer, M. S., Plank, K. M., Rodela, T. M., Sanders, E. R., Schimpf, N. G., Schulte, P. M., Smith, M. K., Stetzer, M., Van Valkenburgh, B., Vinson, E., Weir, L. K., Wendel, P. J., Wheeler, L. B., & Young, A. M. (2018). Anatomy of STEM teaching in North American universities. *Science, 359*(6383), 1468-1470.

Stahnke, R., Schueler, S., & Roesken-Winter, B. (2016). Teachers’ perception, interpretation, and decision-making: A systematic review of empirical mathematics education research. *ZDM Mathematics Education, 48,* 1-27.

Stephen, K. M., & Burns, P. (1986). Teachers' accuracy in self-reporting about instructional practices using a focused self-report inventory. *The Journal of Educational Research, 79*(4), 205-209.

Stephan, M. & Rasmussen, C. (2002). Classroom mathematical practices in differential equations. *Journal of Mathematical Behavior, 21*, 459-490.

Stigler, J. W., Gallimore, R., & Hiebert, J. (2000). Using video surveys to compare classrooms and teaching across cultures: Examples and lessons from the TIMSS video studies. *Educational Psychologist, 35*(2), 87-100.

Stipek, D. J., Givvin, K. B., Salmon, J. M., & MacGyvers, V. L. (2001). Teachers' beliefs and practices related to mathematics instruction. *Teaching and Teacher Education, 17*(2), 213-226.

Streiner, D. L. (2003). Starting at the beginning: An introduction to the coefficient alpha and internal consistency. *Journal of Personality Assessment, 80*(1), 99-103.

Supiano, B. (2018). Are small classes best? It's complicated. *The Chronicle of Higher Education, 64*(29), A20.

Sztajn, P. (2003). Adapting reform ideas in different mathematics classrooms: Beliefs beyond mathematics. *Journal of Mathematics Teacher Education*, *6*, 53-75.

Takeuchi, M. A., & Bryan, V. (2019). Video-mediated interviews to reveal multiple voices in peer collaboration for mathematics learning in groups. *International Journal of Research & Method in Education, 42*(2), 124-136.

Ulam, S. (1976). *Adventures of a mathematician*. New York: Scribner.

U.S. Department of Education. (2006). *A test of leadership: Charting the future of U.S. higher education.* Washington, D.C.: U.S. Department of Education.

Vale, C. D. (1986). Linking item parameters onto a common scale. *Applied Psychological Measurement, 10*(4): 333-344.

Vinner, S. (1991). The role of definitions in the teaching and learning of mathematics. In D. Tall (Ed.), *Advanced Mathematical Thinking* (pp.65-80). Dordrecht, The Netherlands: Kluwer.

Wagner, J. F., Speer, N. M., & Rossa, B. (2007). Beyond mathematical content knowledge: A mathematician's knowledge needed for teaching an inquiry-oriented differential equations course. *Journal of Mathematical Behavior, 26,* 247-266.

Walczyk, J. J., Ramsey, L. L., & Zha, P. (2007). Obstacles to instructional innovation according to college science and mathematics faculty. *Journal of Research in Science Teaching, 44*(1), 85-106.

Walshaw, M. & Anthony, G. (2008). The teacher's role in classroom discourse: A review of recent research in mathematics education. *Review of Educational Research, 78*(3), 516-551.

Ward, R. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association, 58*(301), 236-244.

Wawro, M., Rasmussen, C., Zandieh, M., Sweeney, G., & Larson, C. (2012). An inquiry-oriented approach to span and linear independence: The case of the magic carpet ride sequence. *Problems, Resources, and Innovation in Mathematics Undergraduate Instruction, 22*(8), 557-599.

Wawro, M., Sweeney, G., & Rabin, J. (2011). Subspace in linear algebra: Investigating students' concept images and interactions with the formal definition. *Educational Studies in Mathematics, 78*, 1-19.

Webel, C. & Platt, D. (2015. The role of professional obligations in working to change one's teaching practices. *Teaching and Teacher Education, 47*, 204-217.

Weber, K. (2004). Traditional instruction in advanced mathematics courses: A case study of one professor's lectures and proofs in an introductory real analysis course. *The Journal of Mathematical Behavior, 23*(2), 115-133.

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

Wetzel, D. L. & McDaniel, M. A. (2009). Situational judgement tests: An overview of current research. *Human Resource Management Review,* 19: 188-202.

Wiener, N. (1956). *I am a mathematician.* London: Gollancz.

Wijaya, A., van den Heuvel-Panhuizen, M., & Doorman, M. (2015). Teachers' teaching practices and beliefs regarding context-based tasks and their relation with students' difficulties in solving these tasks. *Mathematics Education Research Journal 27*(4), 637-662.

Wilkins, J. L. M. (2008). The relationship among elementary teachers' content knowledge, attitudes, beliefs, and practices. *Journal of Mathematics Teacher Education 11*(2), 139-164.

Wineburg, S. (1989). Remembrance of theories past. *Educational Researcher, 18*(4), 7-10.

Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement, 73*(6), 913-934.

Wu, H. (1999). The joy of lecturing – With a critique of the romantic tradition of education writing. In S. G. Krantz (Ed.), *How to teach mathematics* (pp. 261-271). Providence, RI: American Mathematical Society.

Yackel, E. & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education, 27*(4), 458-477.

Yackel, E., Cobb, P., & Wood, T. (1991). Small-group interactions as a source of learning opportunities in second-grade mathematics. *Journal for Research in Mathematics Education, 22*(5), 390-408.

Yong, D. (2018, November). *Active learning 2.0: Being intentionally inclusive* [video file]. Retrieved from https://mediaspace.msu.edu/media/t/1_acfszoph .

Yoshinobu, S. (2012, March 22). IBL levels (with apologies to Van Hiele) [Blog post]. Retrieved from http://theiblblog.blogspot.com/search/label/IBL%20Levels.

Yoshinobu, S. & Jones, M. (2012). The coverage issue. *Problems, Resources, and Innovation in Mathematics Undergraduate Instruction,* 22(4), 303-316.

Yoshinobu, S., Jones, M., Cochran, J. J., Cox, L. A., Keskinocak, P., Kharoufeh, J. P., & Smith, J. C. (2011). An overview of Inquiry-Based Learning in mathematics. *Wiley Encyclopedia of Operations Research and Management Science*: John Wiley & Sons, Inc.

Zhu, M., & Ghodsi, A. (2006). Automatic dimensionality selection from the scree plot via the use of profile likelihood. *Computational Statistics & Data Analysis, 51*(2), 918-930.