High School Mathematics Teachers' Perspectives on Selecting, Planning, Setting Up, and Implementing Instructional Tasks With High Cognitive Demand

A dissertation presented to

the faculty of

The Patton College of Education of Ohio University

In partial fulfillment

of the requirements for the degree

Doctor of Philosophy

Otto J. Shaw

August 2021

This dissertation titled

High School Mathematics Teachers' Perspectives on Selecting, Planning, Setting Up, and Implementing Instructional Tasks With High Cognitive Demand

by

OTTO J. SHAW

has been approved for

the Department of Teacher Education

and The Gladys W. and David H. Patton College of Education by

Gregory D. Foley

Morton Professor of Mathematics Education

Sara R. Helfrich

Interim Dean, The Gladys W. and David H. Patton College of Education

Abstract

SHAW, OTTO J., Ph.D., August 2021, Curriculum and Instruction, Mathematics Education

High School Mathematics Teachers' Perspectives on Selecting, Planning, Setting Up, and Implementing Instructional Tasks With High Cognitive Demand

Director of Dissertation: Gregory D. Foley

In school mathematics, students' opportunity to learn varies according to the nature of instruction. Mathematical tasks—that is, problems or activities for student engagement—are critical instructional tools that shape students' mathematical thinking and reasoning. The cognitive demand of a task—the amount, types, and levels of thinking required to solve it—often changes as a teacher modifies the task during planning, setup, and implementation with students. Therefore, school mathematics teachers are instrumental in determining what and how much students learn through their selection and implementation of instructional tasks.

This study explored the perspectives of 9 high school mathematics teachers on their selection, planning, setup, and implementation of mathematical tasks and identified the teachers' reasons for instructional decisions at each of these four phases. Using a thematic analysis approach, the researcher interviewed teachers before and after observing the enactment of a high cognitive demand task. Interviews also focused on teachers' perspectives of how their task unfolded and the cognitive demand associated with each phase of the task. The researcher and a co-observer analyzed each teacher's instructional task as it was (a) selected from curricular source materials, (b) adjusted

during the teacher's planning, (c) set up prior to student engagement, and (d) implemented with students, using the Instructional Quality Assessment rubrics (Boston, 2012) at each phase.

Interview data yielded 18 themes for teachers' task use: 5 for task selection, 5 for task planning, 3 for task setup, and 5 for task implementation. When selecting tasks, teachers frequently considered their learning environment (face-to-face, remote, or hybrid), potential student engagement, real-world contexts, mathematical content, and previous success with the task. During planning, teachers were flexible, adjusted their plan based on the learning environment, attended to recommendations from professional development, established their goals for students, and anticipated difficulties the students might face. The setup phase of instruction typically included a brainstorming stage, a full-class discussion, and teachers' communication of their expectations for student engagement. During task implementation, teachers encouraged productive struggle, asked questions, provided support, facilitated students' engagement, and elicited evidence of students' thinking and reasoning.

The researcher and the co-observer analyzed and reached a consensus on the cognitive demand of each phase of the teachers' tasks by examining documents and observing teachers' instruction. Among the 8 teachers whose tasks were analyzed from selection to implementation, 1 teacher's task increased in cognitive demand from selection to implementation, 3 teachers' tasks maintained cognitive demand, and 4 teachers' tasks experienced a decline in cognitive demand. Interviews with teachers suggested that their perspectives about cognitive demand were sometimes inconsistent with that of the researcher; 3 teachers explained that their task implementation maintained high cognitive demand in instances where the researcher and co-observer's assessment suggested declines in cognitive demand.

The present study suggests that the planning phase is influential in teachers' task use and merits inclusion in the Stein et al. (2009) Mathematical Tasks Framework. Another promising finding is that, of the 8 tasks that were analyzed from selection to implementation, 5 were implemented at high cognitive levels. These teachers had been given curricular materials containing high cognitive demand tasks and were engaged in ongoing professional development to implement the materials effectively. This suggests teacher professional development coupled with cognitively demanding curricular materials as a factor that positively influences teachers' use of instructional tasks, especially when the professional development focuses on the use of high cognitive demand tasks. Based on the findings of this study, future research should further explore teachers' perspectives of mathematical tasks and cognitive demand because mismatches may occur between researchers' and teachers' analysis of the same task. In addition, the methods of the present study should be adapted for use at other grade levels, in other instructional circumstances, and with other content foci.

Dedication

To Ashley, Mom and Dad, and Eli and family.

Without your love and support, I would not be where I am today.

Acknowledgments

I would not have succeeded in earning my PhD without the love and support of my family. Ashley, you have believed in and encouraged me every step of the way; you reassured me when I faltered and encouraged me to keep pushing forward when it got tough. You have worked so hard so that I could further my education, and I cannot thank you enough for all that you are and all that you do.

Mom and Dad, you have always been there for me, especially in these last few years. Thank you for your prayers, guidance, and support that helped me along the way as I worked toward and completed my doctorate. My passion for mathematics education comes from both of you, who taught me the value of lifelong learning and service to others. Eli, Lindsey, Rylee, and Jude, thank you for all the time that Ashley and I have spent with you and for your strong presence in our lives. Family is indeed one of the greatest blessings.

Dr. Greg Foley, thank you for mentoring and advising me these past 4 years. I have learned more about mathematics education than I could have possibly hoped when I was admitted into the program in 2017. My growth as a student, teacher, writer, researcher, and professional are largely thanks to you. You have helped me to go farther than I ever thought possible, and this is only the beginning. To my dissertation committee, Dr. Allyson Hallman-Thrasher, Dr. Mathew Felton-Koestler, Dr. Gordon Brooks, and Dr. Melissa Boston, thank you for your dedication and support throughout my studies and the dissertation process. Each of you substantially contributed to my research and challenged me to produce my best work.

I also want to thank the nine teachers who participated in this study. Amidst the COVID-19 pandemic and on top of your already busy schedules, you made this dissertation possible.

Finally, thank you, God, for Your love and peace which surpasses all human understanding. You are my rock, my foundation, and the anchor that steadies me through every storm. You gave me the strength to keep going when I could not find it within myself. Thank you for opening the door and leading me down the path that I am now on.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Teaching influences student learning. In particular, differing types of instruction provide students with varying opportunities to learn (Boston & Wilhelm, 2017). Instruction focusing on lower-level cognitive processes, such as memorizing facts and performing routinized procedures, enhances students' knowledge and skill in these areas. By contrast, instruction focusing on the higher-level cognitive processes of conceptual understanding and mathematical connections develops students' ability to think and reason mathematically (Jackson et al., 2013; Stigler & Hiebert, 1999). Research has also shown that students' engagement with mathematical tasks influences what and how much they learn (Ni et al., 2018; Stein et al., 2009).

However, few studies have investigated high school mathematics teachers' perspectives as they select, plan, set up, and implement instructional tasks with their students and reflect on these processes. The aim of this study was to explore how high school mathematics teachers use mathematical tasks in their instruction and how they reflect on their decision-making as they select a task from curricular materials, plan to implement the task, set up the task for student engagement, and implement the task with students. Moreover, I sought to identify high school mathematics teachers' perspectives concerning the change in the nature of their tasks throughout this process. By exploring teachers' perceptions, this study intended to enhance researchers', teachers', and teacher educators' knowledge of using mathematical tasks to develop students' mathematical thinking and reasoning.

Background

Improving mathematics teaching and learning has been at the forefront of mathematics education research and reform in the United States for decades (National Council of Teachers of Mathematics [NCTM], 1989, 2000, 2014, 2018; Stein et al., 2009; Stigler & Hiebert, 1999). For example, emphasis on improving school mathematics teaching practices and students' conceptual understanding increased after the Third International Mathematics and Science Study (TIMSS), a video study involving students in Grades 4, 8, and 12 among 41 countries (Stigler & Hiebert, 1999). The TIMSS study in the 1990s indicated that, at the time, U.S. students underperformed in mathematics compared to students in many other countries. It also showed that U.S. mathematics teachers spent the majority of class time focusing on computations and procedures rather than engaging their students in developing conceptual understanding. Stigler and Hiebert asserted that U.S. mathematics teachers participating in the TIMSS study frequently emphasized low-level thinking skills, such as memorization and procedural fluency without connections to underlying mathematical concepts. Mathematics instruction by such teachers typically included demonstrations of how to solve new problems followed by student practice, whereas teachers in higher-performing countries (e.g., Japan) assigned challenging problems and allowed students to develop their own solution pathways.

According to Stigler and Hiebert (1999), "the fact that teaching is a cultural activity explains why teaching has been so resistant to change" (p. 12). Moreover, the National Council of Teachers of Mathematics (NCTM) suggests that improvements in mathematics education have not made a substantial and lasting impact in the United States since the 1996 TIMSS study (NCTM, 2014). The 2007 TIMSS (renamed the Trends in International Mathematics and Science Study) results indicated that, though U.S. middle school students scored just above the international mean in mathematics, they were surpassed by students in Singapore, Korea, and Japan. Moreover, only 6% of U.S. students scored in the "advanced" category, whereas 40% of students in Korea and Singapore and 26% in Japan fell into the same category (Gonzales et al., 2008; Mullis et al., 2008).

U.S. students participating in the Programme for International Student Assessment (PISA) in 2012 struggled with tasks involving real-world models and mathematical reasoning (Organisation for Economic Co-operation and Development [OECD], 2013). Moreover, the ACT 2013 Profile Report and the 2013 SAT Report on College and Career Readiness showed that only 44% of high school graduates were considered ready for college mathematics (ACT, 2013; College Board, 2013). The difficult work of mathematics teachers has only increased since the TIMSS study in the 1990s, as they are faced with "supporting increasingly diverse groups of students to attain increasingly rigorous learning goals" (Cobb et al., 2018, p. 1).

NCTM and other professional organizations have released numerous documents aiming to guide students, teachers, parents, and administrators toward a clearer understanding of the mathematical knowledge required of students and effective teaching practices that support it. In 1989, NCTM's *Curriculum and Evaluation Standards for School Mathematics* presented four curriculum standards for Grades 9–14 relating to mathematical practices: *mathematics as problem solving, mathematics as communication, mathematics as reasoning, and mathematical connections*. In 2000, NCTM released *Principles and Standards for School Mathematics,* emphasizing that the teaching standards should be addressed by focusing on mathematical problem solving and sense making rather than attending to the rote memorization of facts and procedures through extended periods of individual seatwork practicing routine tasks and direct instruction. Instead, NCTM advocates that S*tandards-based* instruction (i.e., instruction tailored to the NCTM *Principles and Standards for School Mathematics*) should focus on the use of mathematical tasks—problems or activities that introduce new mathematical concepts, engage students, and challenge them to think deeply and work hard (Stein et al., 2009; NCTM, 2000).

More recent NCTM documents have focused on inspiring teacher change and action; the statement of teaching and learning standards was no longer enough NCTM. According to *Principles to Actions: Ensuring Mathematical Success for All*, mathematics learning is an active process that requires teachers to challenge, support, and engage students through the implementation of mathematical tasks (NCTM, 2014). To provide students with opportunities to engage in high-level thinking, teachers should frequently select and implement tasks that promote mathematical reasoning and problem solving (NCTM, 2014). In 2018, NCTM published *Catalyzing Change in High School Mathematics: Initiating Critical Conversations* to identify and address challenges in high school mathematics. NCTM (2018) recommends that equitable mathematics teaching includes implementing tasks that allow students to develop positive mathematical dispositions and mathematical identities, in addition to the use of tasks as described in the previous paragraphs.

Throughout the past three decades, NCTM has emphasized the critical importance of instructional tasks and their influence on students' mathematics learning and development. However, the use of tasks does not guarantee that students develop deep, conceptual understandings of mathematics; the teacher's role is influential in determining the mathematics content to highlight, facilitating students' work and discussions, and supporting students without taking over their mathematical thinking and work (NCTM, 2000, 2014; Stein et al., 2009). Students' mathematics learning is therefore influenced by the tasks in which they engage and the actions of teachers to support their engagement.

Problem Statement

In school mathematics, what students learn largely depends upon the tasks they are given and the work they do (Boston & Smith, 2011; Doyle, 1983, 1988). Tasks vary in the types of work and thinking required of students: "tasks that require different cognitive processes are likely to induce different kinds of learning" (Hiebert & Wearne, 1993, p. 395). To support students in developing procedural fluency from conceptual understanding, research suggests that tasks emphasizing such skills should be at the heart of mathematics instruction (Jackson et al., 2013; Silver & Stein, 1996; Stein & Lane, 1996). Such tasks require high levels of cognitive demand, that is, "the kind and level of thinking required of students in order to successfully engage with and solve the task" (Stein et al., 2009, p. 1). Mathematics instruction incorporating high cognitive demand tasks differs from typical U.S. mathematics instruction (Cobb et al., 2018; Stein et al., 2009), and teachers must possess sophisticated knowledge and skill as they respond to what students think, do, and say throughout their engagement with each task. However, empirical evidence suggests that the use of high cognitive demand tasks can enhance students' conceptual understanding, problem-solving skills, and communication skills (Ni et al., 2018; Stein & Lane, 1996; Stigler & Hiebert, 2004), as well as their confidence and dispositions toward mathematics (Boaler & Staples, 2008; Ni et al., 2018; Silver & Stein, 1996).

The complex nature of tasks with high cognitive demand presents teachers with numerous obstacles and barriers to successful implementation. Teachers do not always select tasks based on their cognitive demands (Boston & Smith, 2011; Remillard, 2005); in addition, instructional tasks are not always implemented at their highest potential (Boston & Smith, 2011; Stein et al., 1996). Researchers investigating mathematical tasks have typically focused on teachers' task implementation, the influence of professional development (PD) to support teachers' task implementation, and the relationship between task implementation and student learning. For example, Stein et al. (1996) found that the cognitive demand of a task that students engage with can differ from the cognitive demand of the same task as it appeared in source materials. Henningsen and Stein (1997) identified various factors that tend to influence such change, including classroom management problems and shifting focus to correct answers rather than mathematical understanding. However, more research is needed to investigate teachers' perspectives and rationales as they engage their students in high-level tasks and reflect on their instruction.

Statement of the Research Questions

The present study sought to explore the perspectives of high school mathematics teachers as they select, plan, set up, and implement high cognitive demand tasks. Teachers' reasons for selecting, planning, setting up, and implementing tasks and their perspectives of cognitive demand were of particular interest. Therefore, the research questions guiding this study were the following:

1. When attempting to use a high cognitive demand mathematical task, what actions do high school mathematics teachers take, and for what reasons, while
   a. selecting the task from written source materials?
   b. planning the task for use with their students?
   c. setting up the task immediately prior to student engagement?
   d. implementing the task as students engage with it?
2. What reasons do high school mathematics teachers give to explain the change in the cognitive demand of a task across the four phases of selecting, planning, setting up, and implementing the task?
3. What reasons do high school mathematics teachers give for assessing the cognitive demand of a task at each phase, and in particular, what reasons do they give when there is a mismatch between a teacher's assessment of the cognitive demand of a task and the researchers' assessment of that phase of the same task?

By asking open-ended questions and probing deeper into high school mathematics teachers' selection, planning, set up, and implementation of mathematical tasks, I aimed to uncovered various motivators for their actions and decision making at each stage. The use of open-ended research questions allowed for an inductive approach that captured various factors such as teachers' beliefs and motivation (Beswick, 2011; Collopy, 2003).

The first research question addressed each phase in the progression of a mathematical task. Described in Chapter 2, research by Stein and colleagues (Henningsen & Stein, 1997; Stein et al., 1996, 2009) suggests that mathematical tasks pass through three phases: (a) as they appear in written curricular materials, (b) as they are set up by teachers, and (c) as they are implemented by teachers and enacted by students in the classroom. I have identified and included a fourth task phase, *tasks as they are designed by teachers in their lesson plans,* because teachers may modify a task from how it appeared in source materials (Earnest & Amador, 2019). The tasks that they set up, in turn, may differ from what was planned. Moreover, Jackson et al. (2013) asserted that planning how to engage students when solving mathematical tasks is "central" (p. 655) to setting up tasks that maintain high cognitive demand. As I explain in Chapter 2, there are theoretical (Remillard, 2005) and empirical (James et al., 2016; Smith et al., 2008) bases for including a task planning phase, which was not included by Stein et al. (2009).

The four aforementioned task phases are important to consider because the cognitive demand of a task may differ at each phase, depending on decisions made by the teacher and interactions between the teacher and students (Boston & Smith, 2011; Henningsen & Stein, 1997; Stein et al., 2009). The purpose of the first research question was to identify what high school mathematics teachers do when selecting, planning, setting up, and implementing instructional tasks and why they do these things. As I describe in Chapter 3, the research questions involve tasks that *teachers* perceive as high cognitive demand for two reasons: First, high cognitive demand tasks provide students with greater opportunities to develop conceptual understanding and reasoning skills (Boston & Smith, 2011; Stein et al., 1996, 2009) and are therefore valuable to study. Second, I aimed to explore teachers' perspectives of high cognitive demand tasks and what makes the cognitive demand high.

The second research question focused on (a) the potential change in the cognitive demand of a task between the phases of selecting, planning, setting up, and implementing, and (b) the reasons that teachers gave explain such changes. This question addressed the underlying reasons that teachers associated with the maintenance of high cognitive demand tasks. Previous research suggests that the cognitive demand of a task can decline from one task phase to another. Moreover, high cognitive demand tasks are not always implemented to their highest potential (Boston & Smith, 2011; Cobb et al., 2018; Henningsen & Stein, 1997). Considering these findings, the second research question explored high school mathematics teachers' perspectives considering changes in cognitive demand from one task phase to the next. For example, one factor that tends to support high level task implementation in the literature is appropriate teacher scaffolding (Henningsen & Stein, 1997; Stein et al., 2009); additional research is needed illuminate teachers' decisions about when to provide scaffolding and how much support to provide during task implementation. Examining the underlying reasons behind such decisions has provided further insights to support high school mathematics teachers' implementation of high cognitive demand tasks.

The third research question explored how teachers' analysis of a task compared to that of the researcher. Determining the cognitive demand of a task can be challenging because low-level tasks often possess surface-level features consistent with reform-oriented tasks (Boston & Smith, 2011; Stein et al., 2009). For example, low-cognitive tasks involving the use of manipulatives, real-world contexts, or diagrams may appear to be high in cognitive demand; however, such tasks that imply the use of well-rehearsed procedures are considered to be low in cognitive demand by mathematics educators and researchers (e.g., Stein et al., 2009).

As I explain in Chapter 3, I required research participants to select, plan, set up, and implement what they considered to be high cognitive demand tasks and had them analyze the tasks using the Stein et al. (2009) Task Analysis Guide (TAG), a tool designed for teachers to learn about and analyze instructional tasks. Instead, I used the Instructional Quality Assessment (IQA) toolkit (Boston, 2012), an instrument that was influenced by the TAG but requires training to implement as intended by its developers. In instances where a mismatch occurred (i.e., a teacher identified a task as in high cognitive demand using the TAG, but I identified it as low in cognitive demand using the IQA), I sought to determine potential reasons why such mismatched might have occurred. In Chapter 3, I explain that I did not directly probe teachers for additional information when mismatches occurred because I did not want to influence their analysis of their tasks. Instead, I considered ways in which our analyses of the task differed.

Research Purpose

The goal of this study was to explore the perspectives of high school mathematics teachers' as they engaged in selecting, planning, setting up, and implementing instructional tasks. As past studies have examined the use of mathematical tasks from a behavioral perspective (e.g., teacher actions that lower or maintain cognitive demands), this study delved deeper and investigated the underlying reasons influencing these decisions. This study illuminated how high school mathematics teachers conceptualize the process of selecting, planning, setting up, and implementing mathematical tasks. Additionally, I have identified various reasons that high school mathematics teachers attribute to the high cognitive demand of tasks and the potential changes in cognitive demand between task phases.

The primary focus of this research was to investigate how high school mathematics teachers use mathematical tasks and what guided their decision-making. Specifically, I was interested in the use of *high cognitive demand* tasks because they can promote students' conceptual understanding, problem-solving skills, and communication skills (Boston & Smith, 2011; Hiebert & Wearne, 1993; NCTM, 2018), aligning with the standards and recommendations of mathematics teaching reform efforts. Identifying the underlying reasons for teachers' task use has provided evidence of effective teacher mindsets and rationales that may enhance the practice of high school mathematics teachers and lead to improvements in student learning.

Henningsen and Stein (1997) and Stein et al. (1996), among others, have identified reasons for the potential change in the cognitive demand between setup and implementation. These factors are shown in Table 1. For example, the selection of a mathematical task that builds on students' prior knowledge is based on the teacher's action of selecting such task. However, previous research has not explored teachers' perceptions and the reasons underlying such actions. Research investigating *why* and *how* a task was selected has provided insights that can enhance mathematics teachers' ability to analyze tasks based on their cognitive demand and select tasks accordingly. This study built on previous research by exploring some of the Henningsen and Stein (1997) factors from teachers' perspectives and revealing the underlying reasons associated with such factors.

Table 1

Factors Influencing the Maintenance or Decline of Cognitive Demand

Cognitive Demand Is Maintained	Cognitive Demand Declines
Tasks build on students' prior knowledge	Inappropriateness of the task
Teacher scaffolding	Classroom management problems
Appropriate amount of time	Too much or too little time
High-level performance modeled	Lack of accountability
Teacher presses for explanations	Challenges become nonproblems
Teacher draws conceptual connections	Focus shifts to correct answer

Note. This table includes factors identified by Henningsen and Stein (1997) that influence changes between task selection and implementation.

Educational Significance

Examining high school mathematics teachers' actions and reasons for selecting, planning, setting up, and implementing tasks has several implications. The impetus for this research was to understand what makes high school mathematics teachers effective when selecting, planning, setting up, and implementing high cognitive demand tasks because such tasks can be beneficial for students' conceptual understanding (Stein et al., 2009; NCTM, 2014). Moreover, high cognitive demand tasks can enhance students' mathematical understanding and skills consistent with reform efforts in mathematics education. For example, The National Research Council (NRC) defines *mathematics proficiency* as comprised of five elements: conceptual understanding, procedural fluency, strategic competence, adaptive reasoning, and productive disposition (NRC, 2001). To help students meet these goals, NCTM recommends the use of high-leverage teaching practices (Ball & Forzani, 2010), such as implementing tasks that promote reasoning and problem solving, facilitating meaningful mathematical discourse, and building fluency from conceptual understanding (NCTM, 2014). High cognitive demand tasks can help students reach these goals, when implemented using effective instructional practices (Boston & Smith, 2011; NCTM, 2014).

The *Common Core State Standards for Mathematics* (CCSSM) include standards for mathematical practice that should be met by *all* high school students (National Governors Association Center for Best Practices and Council of Chief State School Officers [NGA & CCSSO], 2010). These standards include "make sense of problems and persevere in solving them," "reason abstractly and quantitatively," and "use appropriate tools strategically" (pp. 6–7). The traditional approach of having students memorize facts and apply procedures mindlessly does not help students to meet these standards. However, mathematical tasks that are planned, set up, and implemented at high cognitive levels can help students develop mathematics proficiency and meet the CCSSM standards for mathematical practice (NCTM, 2014). Evidence of teachers' perspectives, including the successes and challenges of high-level task implementation, contributes to efforts that support teachers in this work. Through this study, teachers were given opportunities to explain how they made instructional decisions, providing insights that have illuminated additional obstacles to high-level task implementation and methods of overcoming them.

This study has benefited participating teachers by challenging them to think more deeply about their instruction and how it impacts student learning. By reflecting on the ways that they select, plan, set up, and implement mathematical tasks, research participants have enhanced their task analysis skills which will benefit their future practice. Interviews with research participants focusing on the underlying reasons that impact their instructional decisions have enhanced participants' awareness of such reasons and decisions; increased awareness may lead these teachers to challenge their preexisting notions and reflect upon how these notions influence their task use. For example, beliefs about mathematics curriculum may guide teachers' decisions concerning which tasks to select and how to modify them during their planning (Carson, 2011; Collopy, 2003; Philipp, 2007). Helping teachers to become more aware of these beliefs may evoke change that positively impacts how they analyze tasks.

This study has implications for mathematics teacher education. Evidence of teachers' perspectives regarding mathematical tasks may help to guide mathematics teacher educators' interactions with preservice teachers. Preservice mathematics teachers are likely to demonstrate some elements of similar perspectives, beliefs, attitudes, and rationales as the participants of this study; therefore, this study may serve to guide mathematics teacher educators as they instruct and guide their teacher candidates. This study may also guide future PD programs, especially those emphasizing the use of high cognitive demand tasks, by identifying ways to support teachers as they learn to use mathematical tasks in their instruction. Identifying teachers' reasons for high-level task use at each task phase and reasons for change in cognitive demand will provide future PD programs and preservice teacher education programs with goals and strategies that support effective task use. Moreover, the present study may provide insights on the perspectives of high school mathematics teachers who choose to participate in task-based PD in the future.

Finally, this study has addressed a gap in the literature concerning mathematics teachers' use of instructional tasks. Tasks as they appear in curricular materials tend to differ from the tasks implemented by teachers (Boston & Smith, 2011; Stein et al., 1996, 2007). Research has also identified various factors contributing to differences between tasks as set up and as implemented (Henningsen & Stein, 1997). However, less is known about what influences teachers' decisions when selecting tasks, their rationales for modifying source tasks for their lesson plans, and their reasons for modifying tasks further when setting them up for students. Research investigating mathematics teachers' beliefs (e.g., Philipp, 2007) has suggested implications for mathematical tasks at each phase, but empirical evidence is lacking. The present study addressed this gap in the literature by identifying reasons that motivate the factors identified by Henningsen and Stein and reasons for teachers' decision-making between the other task phases.

Delimitations

The present study was delimited to the selection of high school mathematics teachers in the state of Ohio who have engaged in PD emphasizing the use of high cognitive demand tasks. Selecting teachers who have experience and training with the TAG allowed for the recruitment of participants who were more likely to select, plan, set up, and implement mathematical tasks at high cognitive levels. Such teachers were also more likely to be familiar with the constructs and language such as *mathematical tasks* and *cognitive demand*. Moreover, interviews investigating teachers' perspectives captured meaningful and useful data because the teachers were already somewhat familiar with the Stein et al. (2009) Mathematical Tasks Framework (MTF) and TAG and the cognitive demands of tasks they used were initially high, as discussed in Chapter 4.

The purposeful sample for this study was taken from high school mathematics teachers in the state of Ohio for several reasons. As I describe in Chapter 3, recent PD programs in Ohio have focused on implementing instructional change at the high school level: two of such PD programs are for (a) the Ohio Mathematical Modeling and Reasoning (MMR) pilot course developed by the Ohio Department of Education (ODE) and (b) Advanced Teacher Capacity (ATC) programs hosted at Ohio University. The Ohio MMR course was pilot tested for its third year during the 2020–2021 academic year, providing a source of research participants who had experienced relevant PD. Similarly, the two PD programs offered through ATC served high school mathematics teachers and incorporated the Stein et al. (2009) MTF and TAG as integral PD components. Teachers involved in the MMR and ATC programs were also located within a close proximity of Ohio University; therefore, some of them have developed professional relationships with Ohio University faculty and graduate students, which provided me further accessibility and rapport with those who were willing to engage in research. Though remote data collection procedures resulting from the COVID-19 pandemic allowed for the recruitment of teachers outside Ohio, I kept within these delimitations for the aforementioned reasons.

Limitations

This study was limited in the sense that, as qualitative research, results may not necessarily generalize to the entire population of high school mathematics teachers in Ohio (Glesne, 2016; Patton, 2015). Rather, this dissertation reports descriptive results pertaining specifically to the participating teachers involved. However, the themes emerging from the data may be transferrable to other contexts and situations (Lincoln & Guba, 1985; Merriam, 2009). Additionally, this study was limited to my own data collection and analysis abilities, as I was the instrument of this qualitative study (Merriam, 2009; Patton, 2015). The data for this study were collected and analyzed by myself, limited to my positionality and capabilities as a researcher, and furthermore, the analyses were conducted based on my own knowledge and abilities. However, an additional trained IQA rater observed instruction and coded tasks using the rubrics to provide reliability and credibility to the results obtained from observations and analysis of student work (Boston, 2012; Glesne, 2016).

Another limitation resulted from the COVID-19 pandemic that swept the world in 2020 and 2021. The disease COVID-19 severely impacted U.S. schools because of its potential to cause significant health issues and even death among individuals. To prevent the spread of the virus and keep students and school personnel safe, many Ohio districts transitioned from traditional face-to-face instruction to online learning formats; others elected to remain face-to-face with various procedures in place to protect individuals' health (e.g., the use of face masks, social distancing, and barriers around desks), whereas some districts implemented hybrid formats (i.e., some students learned face-to-face, and others learned remotely). As I describe in Chapters 2 and 3, the data collection procedures for the present study were influenced by the COVID-19 pandemic because the data were gathered in the fall of 2020 and the spring of 2021. For example, all interviews and observations of teachers' instruction were done remotely via Zoom to protect the health of research participants, students, and the researcher. Adjustments due to the COVID-19 pandemic introduced limitations to the study because I was unable to visit teachers' classrooms and conduct interviews in person. However, I conducted the present study as rigorously and ethically as possible despite the unforeseen setbacks.

Definitions of Key Terms

Mathematical task refers to a problem or activity that is used to engage students during mathematics instruction (Boston & Smith, 2009; Stein et al., 1996).

Mathematical task as it appears in source materials may be a mathematical task that is written in a textbook, found online, or presented in another source (Stein et al., 1996, 2009). This includes any printed or digital text, images, figures, and tables provided in the statement of the task.

Mathematical task as planned refers to a mathematical task as it appears in teachers' written lesson plans, including representation(s) of a task and any instructional notes included in the lesson plan. This also refers to the print or digital version of a task that a teacher modifies for their instruction in any way. Though Stein et al. (2009) do not include this task phase, I include teachers' planning because teachers may modify tasks between this phase and the setup phase (Earnest & Amador, 2019). For example, a teacher may choose to provide more explicit directions based on how a task unfolded in a previous period. Such instructions might not have been included in the original lesson plan.

Mathematical task as set up refers to a mathematical task as it is initially presented to students in class, including teacher directions given orally and in writing (Stein et al., 1996, 2009). Task setup may also include the distribution of materials and tools for student use and discussions of what is expected of them. The setup phase occurs immediately prior to implementation phase, during which students engage in solving the given task. Based on the work of Jackson and colleagues (2013), this phase of instruction also includes whole-class discussions that occur prior to student work time. Such discussions may attend to the contextual features or the mathematical relationships of a task that are necessary for students to engage with it effectively.

Mathematical task as implemented refers to a mathematical task as implemented by the teacher and enacted by students. This includes any verbal and written communication made by the teacher and students while solving the task and all work that is done to complete the task immediately following task setup (Stein et al., 1996, 2009). As described in Chapter 3, I conducted observations to obtain evidence of interactions between teachers and students and I collected student work samples to measure the quality of the work done by students. Tasks at this phase are enacted by both the students and the teacher (Boston & Smith, 2011; Stein et al., 2009), but I use the term *implemented* to emphasize the teacher's actions, the focus of the present study.

Cognitive demand is "the kind and level of thinking required of students in order to successfully engage with and solve a task" (Stein et al., 2009, p. 1). Cognitive demand can be measured using the Task Analysis Guide and has four levels: *memorization*, *procedures without connections*, *procedures with connections*, and *doing mathematics* (Stein et al., 2009), each of which is described in Chapter 2. Cognitive demand can also be measured using the Instructional Quality Assessment, including levels of 1–4 which closely map onto the four TAG levels (i.e., a score of 1 closely represents *memorization*, a score of 2 closely represents *procedures without connections*, and so on), though levels 3 and 4 are both used to categorize *procedures with connections* and *doing mathematics* tasks; the factor that distinguishes a score of 3 from a score of 4 is the explicit prompt for evidence of students' thinking and reasoning. A score of 0 in the IQA indicates nonmathematical activity, the lowest level that is not included in the TAG explicitly

(Boston, 2012). Further description of the IQA and each score level are provided in Chapters 2 and 3.

High cognitive demand task (cognitively demanding task) is a mathematical task at any task phase which is as the level of *procedures with connections* or *doing mathematics* according to the Stein et al. (2009) TAG or a task at any phase with an IQA score of 3 or 4 (Boston, 2012). Such tasks require higher cognitive demand than tasks at the level of nonmathematical activity, *memorization*, and *procedures without connections*.

Outline of the Study

In this chapter, I introduced the research study, including the background, problem statement, research questions, research purpose, significance of the study, delimitations, limitations, and definition of key terms. Chapter 2 is a review of the literature, including research on mathematical tasks and cognitive demand, mathematics teacher beliefs, the influence of PD on teacher change, and the IQA. In Chapter 3, I discuss my research methods, including the research design, context of the study, the selection of research participants, and data collection and analysis procedures. In Chapter 4, I report the findings of the study, including analysis of individual cases (teachers and their tasks) and an analysis of themes and trends across cases. In Chapter 5, I discuss these research findings, their relationship to the literature, and implications for further research.

Chapter 2: Literature Review

This review of the literature is organized into three sections: (a) research involving mathematical tasks; (b) the influence of knowledge, beliefs, and professional development (PD) on how high school mathematics teachers select, plan, set up, and implement mathematical tasks; and (c) the development and use of the Instructional Quality Assessment (IQA) Classroom Observation Toolkit. Mathematical task research is divided into three subsections:

- research on academic tasks and academic work,
- research on mathematical tasks and cognitive demand, stemming from the work on academic tasks and academic work, and
- empirical studies investigating the use of mathematical tasks.

The section mathematics teachers' knowledge, beliefs, PD, and teacher change is divided into four sections:

- definitions of knowledge and beliefs,
- mathematics teacher knowledge and beliefs and their impact on selecting, planning, setting up, and implementing mathematical tasks,
- the development of a conceptual framework based on the previous subsections, and
- a model for PD and teacher change that incorporates teacher knowledge and beliefs.

The section describing the IQA consists of

- a discussion of the theoretical frameworks underlying the IQA rubrics,

- a brief report on empirical studies using the IQA, including research testing validity and reliability, and
- the Expanded IQA Task Setup Rubrics.

Mathematical Tasks

The following subsections include relevant literature pertaining to mathematical tasks, including the history of tasks as a theoretical construct and the development of the construct over time. Doyle's (1983, 1988) synthesis of literature brought about the idea of academic tasks and academic work, leading to the development of Stein et al.'s (1996) definition of mathematical tasks and cognitive demand. Moreover, Doyle's classification of academic tasks and the varying levels of cognitive demand required for students to solve tasks influenced the Task Analysis Guide (TAG) developed by Stein and colleagues. This section concludes with a description of empirical studies investigating teachers' use of mathematical tasks, emphasizing the lack of research exploring mathematics teachers' thought processes and rationales for the decisions they make when selecting, planning, setting up, and implementing mathematical tasks.

Academic Tasks and Work

The foundation for mathematical tasks research began with Doyle (1983, 1988), who aimed to investigate the "missing element" (Doyle, 1988, p. 168) in the study of academic work at the time. Though numerous researchers had studied the types and amount of work done by students in school (e.g., Borg, 1980; Johnson, 1980; Rosenshine, 1980). Doyle (1983) argued that these studies provided "little sense of the inherent demands of that work" (p. 160), motivating the development of *academic tasks* as a research construct. The term *task* highlights four aspects of students' work: (a) the products students are to create, (b) the processes used to create the product, (c) the resources available to students as they create the product, and (d) the weight or significance of a task, meaning the weight assigned to assignments, quizzes, and tests in determining students' grades (Doyle, 1988). For example, a product might be the answer to a mathematical question, the processes used might be the algorithms, formulas, and procedures used, and an available resource might be an example of a similar problem worked out by either the teacher or another student. Academic tasks are thusly defined as "the answers students are required to produce and the routes that can be used to obtain these answers" (Doyle, 1988, p. 161). Teachers influence tasks, and therefore student learning, by identifying and shaping the work students do in class (Doyle, 1983, 1988).

A task is a basic unit of curriculum. It serves to identify content and processes for student focus. Tasks are useful as a tool to analyze instruction because they define the work students do in class, influence their learning, and determine how they think about a subject and its meaning (Doyle, 1983, 1988). Academic tasks, however, require various cognitive processes to solve. Doyle defines four kinds of academic tasks based on the type and amount of student cognition required:

1. *memory tasks*, requiring the direct recall of known information,
2. *procedural or routine tasks*, requiring the application of a known or predictable algorithm to solve,
3. *comprehension or understanding tasks*, in which students must apply known information or procedures to new contexts and situations, and

4. *opinion tasks*, requiring students to give their opinion for something.

The cognitive processes required to solve the first three task types increases from one to the next, as procedural tasks apply what is already known (without necessarily understanding the underlying meaning) and comprehension tasks apply potentially familiar procedures to new contexts (typically requiring an understanding of the underlying meaning of the procedure and when it is applicable). “A procedural task is one that can be accomplished without understanding by simply knowing how to follow a series of computational steps. Understanding tasks, on the other hand, requires knowledge about why the computational steps work” (Doyle, 1988, p. 165).

The nature of a task may also vary depending on its “level” (Doyle, 1988, p. 170): when announced by the teacher, when interpreted by students, and when reflected in the work done by the students and accepted by the teacher. Though curriculum guides tend to emphasize the use of higher cognitive processes involved in comprehension or understanding tasks, according to Doyle, low-cognitive tasks (memory and procedural tasks) are common in classrooms. Rather than analyzing and strategically implementing tasks with students, mathematics teachers typically “focus instruction on computational procedures and accuracy of calculations” (p. 171). Within this mode of instruction, students generally know which procedures or algorithms they must use to solve problems in advance (Doyle, 1988; Stigler & Hiebert, 1999).

Academic tasks do not exist in isolation; rather, they become increasingly complex when connected to the classroom setting, possibly explaining the common use of low cognitive tasks in mathematics instruction. From the perspective of teachers, academic tasks are implemented in a classroom where organizing and facilitating students' engagement requires effective classroom management skills and complex social interactions with students. Doyle (1983) cites evidence that the inability to organize and manage classroom environments can negatively influence academic work, and thusly, student achievement (Brophy, 1979; Good, 1979). For students, academic work is situated in the evaluative nature of school, introducing both ambiguity (the extent to which an answer to an academic task can be defined in advance) and risk (the strictness of evaluative criteria used by the teacher and the ability to meet these criteria) as "inherent features" of academic work (Doyle, 1983, p. 183). Ambiguity and risk vary from one academic task to the next and are not necessarily dependent on one another; that is, a considerably ambiguous task might not contain much risk and vice-versa. Doyle categorizes the four task types defined previously based on the ambiguity and risk associated with each:

1. Memory I and Routine I tasks, low in both ambiguity and risk,
2. Memory II and Routine II tasks, high in risk (larger tasks, requiring more work to be done than Memory I and Routine I) but low in ambiguity,
3. Opinion tasks, low in risk but high in ambiguity, and
4. Understanding tasks, high in both ambiguity and risk.

Tasks with higher ambiguity and risk are more difficult to implement with students for various reasons. Doyle (1983) argued that students sometimes provide obstacles, inventing their own strategies to reduce ambiguity and risk. Students may minimalize their responses to academic tasks (Graves, 1975; Rosswork, 1977) and devise strategies to put the required work onto others, sometimes hesitating to respond until another student or the teacher answers a question for them (MacKay, 1978; Mehan, 1974). In mathematics classes, Doyle (1983) provided evidence that students resist shifts from routine and procedural tasks to understanding tasks by refusing to cooperate until they are explicitly told what to do (Davis & McKnight, 1976). When faced with these situations, it is not surprising that teachers shift from understanding and high-level cognitive processes toward memory and routine tasks, which “substantially” (Doyle, 1983, p. 186) reduce the complexity of classroom management issues. “The type of tasks which cognitive psychology suggests will have the greatest long-term consequences for improving the quality of academic work are precisely those which are the most difficult to install in classrooms” (p. 186). Teachers face pressure to reduce the use of understanding tasks and focus on helping students get their work done, rather than focusing on the quality of the work (Doyle, 1983; 1988).

Doyle’s (1988) concept of academic work in mathematics class, a function of the academic tasks in which students engage, is divided into two categories: *familiar work* and *novel work*. Familiar work tends to incorporate memory and procedural tasks, whereas novel work, consisting of understanding tasks, requires students to synthesize relevant information and apply knowledge in unfamiliar contexts. Novel work entails that “students must make decisions about what to produce and how to produce it” (Doyle, 1988, p. 173). Therefore, novel tasks contain a considerable amount of ambiguity and risk due to their unpredictability and high cognitive demands.

Doyle (1988) observed middle school mathematics classes that typically aimed for the large accomplishment of academic work using "routinized work patterns" (p. 175), including explicit demonstrations of how to solve problems and extensive guided practice. Students were "seldom required to assemble information or processes in ways that had not been demonstrated to them in advance" (p. 175). Mathematics curriculum in these classrooms consisted of independent, unconnected skills emphasizing computational accuracy rather than conceptual understanding and problem solving (Doyle, 1988). Doyle's work suggests that the consistent emphasis on familiar work and low cognitive tasks may not only be due to the difficulty implementing high level tasks, but also because standardized tests primarily consist of memory and procedural tasks; familiar work is seen as appropriate for handling complex classroom environments and preparing students for the work they must do to perform on standardized tests.

Phases in the Life of a Mathematical Task

Mathematical tasks, extending from Doyle's (1983, 1988) notion of academic tasks, are problems or activities that are used to engage students during mathematics instruction and serve as influential factors that determine what and how much students learn (Boston & Smith, 2011; Stein et al., 1996, 2009). Mathematical tasks are dynamic in the sense that they progress through several stages before they are experienced by students. According to the Mathematical Tasks Framework (MTF) (Stein et al., 2009), tasks move through three phases: first, as they appear in curricular or instructional materials; second, as they are set up by teachers; and third, as they are implemented by students. The MTF depicts that student learning follows from the tasks that are implemented by students; however, each of the task phases influence student learning in some way (Boston & Smith, 2011; Stein et al., 2009).

Distinguishing between the three MTF task phases is necessary because mathematical tasks can potentially change between each phase (Ni et al., 2018; Stein et al., 2009). Stein et al. (1996) defined task setup as:

> the task that is announced by the teacher. It can be quite elaborate, including verbal directions, distribution of various materials and tools, and lengthy discussions of what is expected. Task set up can also be as short and simple as telling the students to begin work on a set of problems displayed on the blackboard. (p. 460)

However, the task that a teacher sets up may not always be identical to the original task as it appeared in written source materials. For example, a teacher may choose to include, remove, or modify elements of a task based on past experiences and goals for teaching (Earnest & Amador, 2019; Grouws et al., 2013).

Remillard (2005) asserted that “the curriculum enacted in the classroom can, at best, be represented by the curriculum planned by the teacher” (p. 238). The importance of teacher planning for mathematical task setup and implementation was made evident by Smith et al. (2008), who described a framework for developing lessons called the Thinking Through a Lesson Protocol (TTLP). The TTLP is a tool that helps teachers to anticipate students’ thinking and work during a lesson by posing questions for teachers to consider as they select and set up a task, support students’ exploration, and facilitate discussions (Smith et al., 2008). Through the Justification and Argumentation: Growing

Understanding of Algebraic Reasoning (JAGUAR) project, a team of researchers and middle school teachers found that using the TTLP helped teachers to both plan and implement mathematical tasks rigorously (James et al., 2016). By having teachers work through the mathematical tasks they anticipated implementing with their students and "doing the math" (p. 417) in advance, James et al. found that teachers were able to identify the potential of a task, the goals they had for students, the prior knowledge needed to work on the task, and possible solutions students might come up with. Though not identified by Stein and colleagues (2009) in the MTF, *tasks as they are designed by teachers in their lesson plans* is a critical task phase that stands between task selection and task setup.

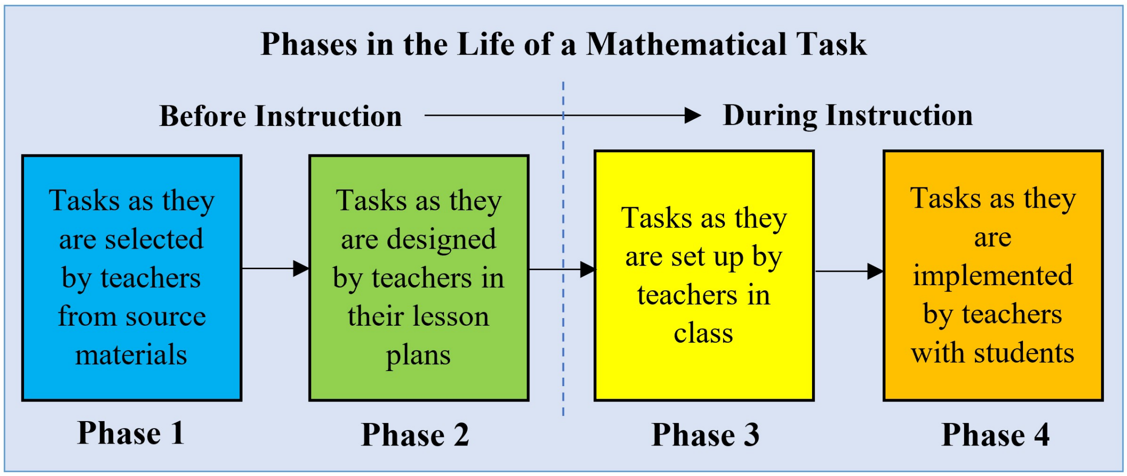
Task implementation has been defined as the "manner in which students actually work on the task" (Stein et al., 1996, p. 460). Just as teachers may modify tasks during planning, students might not necessarily work through a mathematical task in the way that it was set up for a variety of reasons (e.g., teachers might choose to simplify tasks because they feel the task is beyond their students' mathematical understanding). Stein et al. (2007) examined the nature of task changes from a curricular standpoint, highlighting that research on teaching and curriculum shows substantial differences between tasks as they appear in source materials and tasks as implemented with students. The enacted curriculum (tasks as they unfold in the classroom) and the experienced curriculum (what is learned by students from engaging in the tasks) are influenced by interactions between teachers and students, teachers' beliefs and knowledge, teachers' professional identities, and other factors, creating something that differs from the written task and the plans made by the teacher (Earnest & Amador, 2019; Stein et al., 2007).

Figure 1 depicts four phases of a mathematics teacher's task use, including three phases from the MTF plus a planning phase between task selection and task setup. Phase 1: Selection, *tasks as they are selected by teachers from source materials*, is similar to the first phase in the MTF—it intends to capture tasks that appear in written curricular materials (e.g., textbooks or the internet). Phase 2: Planning, *tasks as they are designed by teachers in their lesson plans*, was described in the previous paragraphs and was not included in the MTF. In this phase, "lesson plans" refer not only to actual written plans but also to any modifications made by teachers to the task from the first phase, because not all teachers create and use written lesson plans.

In Figure 1, Phases 3 and 4, *tasks as they are set up by teachers in class* and *tasks as they are implemented by teachers with students*, respectively, are comparable to the setup and implementation phases identified in the MTF. However, I designed Figure 1 to include language focusing on teachers' actions (selecting, planning, setting up, and implementing tasks) because the purpose of the present study is to explore teachers' perspectives and how they inform their actions. Consequently, I dropped *student learning,* which was included in the MTF. In Figure 1, the vertical line segment dividing the first two and the last two phases separates the teacher actions that occur prior to instruction from those that occur during instruction. This separation is also important for data collection procedures, discussed in Chapter 3.

Figure 1

The Four Phases of Mathematics Teachers' Task Use



Note. This figure is based on the Mathematical Tasks Framework by Stein et al. (1996, 2009) but emphasizes the actions of teachers at each task phase, omits the student-learning phase, and adds a teacher-planning phase (Phase 2) not included in the MTF.

The Cognitive Demand of a Mathematical Task

Stein and colleagues (2009) emphasized that "students need opportunities on a regular basis to engage with tasks that lead to deeper, more generative understandings regarding the nature of mathematical processes, concepts, and relationships" (p. 5). Such tasks require high levels of cognitive demand, meaning that the tasks require complex thinking and reasoning from students to solve. Stein et al. (2009) defined four categories of mathematical tasks based on the cognitive demands required of students to complete them, as represented in the Task Analysis Guide (TAG): *memorization, procedures without connections, procedures with connections,* and *doing mathematics*:

1. *Memorization* tasks, similar to Doyle's (1983) "memory tasks" (p. 162), are straight forward, requiring students to recall factual information.
2. *Procedures without connections* tasks, analogous to Doyle's "procedural or routine tasks" (p. 163), are algorithmic, but have no connections to underlying concepts or meaning.
3. *Procedures with connections* tasks capture some of the essence of Doyle's (1988) "comprehension or understanding tasks" (p. 163); these tasks differ from *procedures without connections* in that they require students to connect procedural fluency and conceptual understanding.
4. By *doing mathematics*, the authors refer to tasks that engage students in exploring mathematical concepts, processes, and relationships (Stein et al., 2009).

The cognitive demand of a mathematical task can be analyzed during each phase either as shown in the MTF or as shown in Figure 1. The cognitive demand can—and often does—change from one phase to the next.

Consider the following example of a mathematical task for students in Grades 6–8, presented by Stein and Smith (1998): "What are the decimal and percent equivalents for the fractions 1/2 and 1/4?" (p. 269). According to Stein and Smith, an expected student response might be that 1/2 = 0.5 = 50% and 1/4 = 0.25 = 25%, making this a *memorization* task. Students are simply asked to state the required information without providing any pictorial or written explanation for their answer. For this task, either students already know the answer, or they do not.

However, consider another task addressing the same mathematical content from a different perspective:

> Shade 6 small squares in a 4 × 10 rectangle. Using the rectangle, explain how to determine each of the following: (a) the percent of area that is shaded, (b) the decimal part of the area that is shaded, and (c) the fractional part of the area shaded. (Stein & Smith, 1998, p. 269)

Stein and Smith identify this as a *doing mathematics* task because students are not explicitly told which procedure(s) to use and are asked to explore the relationships between mathematical representations. Students are also held accountable for providing explanations for each of the three representations (percent, decimal, fraction) they are asked to identify, with relation to the shaded rectangle provided. Each of these mathematical tasks, at this point, can only be assessed from a written perspective; planning and implementation of these tasks may vary and require separate analysis.

Maintaining the Cognitive Demand of a Task

To engage their students in developing rich, conceptual understandings of mathematical concepts, mathematics teachers should begin by selecting cognitively demanding tasks to use during instruction. However, research shows that "teachers typically do not analyze tasks in terms of cognitive demands" (Boston & Smith, 2009, p. 123). Rather, teachers tend to select tasks based on "surface-level features" (p. 123), such as whether a task is given as a word problem or whether a task involves graphing. Teachers often turn to their textbooks as a primary source of instructional tasks (Remillard, 2005) and select tasks that focus on specific skills, concepts, or standards they wish to address (Boston & Smith, 2011). As shown in the MTF and in Figure 1, the selection of a cognitively demanding task is followed by additional obstacles: planning to engage students in the task, setting up the task, and implementing it with students while maintaining the mathematical and cognitive rigor.

The way in which a mathematical task is presented by the teacher directly influences how students will engage with and complete the task, which, in turn, determines the nature of students' opportunities to learn. For example, teacher questioning in both the setup of the task and during implementation is an essential component of instruction that determines whether cognitive demands are maintained (Ni et al., 2018; Stein et al., 2007). Challenges faced by teachers include difficulty with giving control and authority to students and knowing when to ask questions versus allowing students to struggle. It then becomes tempting to simplify the task by explicitly telling students which procedures or steps to use. However, doing so eliminates the need for students to work and think about how they should approach the task, and ultimately lowers the cognitive demand required to solve it (Boston & Smith, 2011; Stein et al., 2009).

The cognitive demand of a task can also lower when teachers fail to hold their students accountable for high quality work (Boston & Smith, 2011; Stein et al., 2009). When announcing a task, teachers can avoid this by clearly stating that students must explain their thinking and justify their work (either orally or in writing). Moreover, effective normative practices for engaging students in high-level tasks are scaffolding student thinking, modeling high-quality work and thinking, pressing for explanations and justifications, and selecting tasks that build on students' prior knowledge (Stein et al., 2007). Teachers' beliefs regarding their students' mathematical ability can also negatively impact the enacted and experienced curriculum of mathematical tasks. According to Stein et al. (2007),

> Teachers tend to be concerned about the high level of independent thought, problem solving, and self-monitoring demanded of students by the tasks found in standards-based curricula, expect that students cannot manage these demands, and consequently restructure or adapt the lessons to make them less complex and more readily accessible to students. (p. 355)

In doing so, teachers limit the opportunities for students to engage in high-level thinking, reasoning, problem solving, and communication skills afforded by high-quality curricular materials.

Mathematical tasks that are designed as *procedures with connections* and *doing mathematics* have the potential to be implemented in less cognitively demanding ways depending on how the teacher sets up the task and how students engage in it (Boston & Smith, 2009, 2011; Stein et al., 2009). However, research shows that teacher training can be effective in encouraging teachers to select and maintain cognitively demanding tasks. For example, Boston and Smith (2011) found that teachers who participated in a task-focused PD program "improved their ability to select and implement cognitively demanding tasks during the time frame of the professional development initiative and sustained this ability a year after their involvement in the project ended" (p. 974). Moreover, a study by Arbaugh and Brown (2005) showed that engaging teachers in PD related to the (Stein et al., 2009) TAG encouraged teachers to think more deeply about the relationship between mathematical tasks and the work of their students in a nonthreatening way. Mathematics teachers who examine and reflect on their practice, considering the cognitive demand of the tasks they select and their implementation, can begin to improve their practice by getting their students to think at higher levels.

Empirical Studies Investigating the Use of Mathematical Tasks

In 1990, the University of Pittsburg initiated the QUASAR (Quantitative Understanding: Amplifying Student Achievement and Reasoning) Project, a 5-yr project investigating mathematics instructional programs for middle school students in economically disadvantaged areas (Silver & Stein, 1996; Stein et al., 1996). Stein and colleagues examined the consistency between mathematical tasks as set up and as implemented, with emphasis on solution strategies possible, mathematical representations possible (e.g., algebraic equations, graphs, numerical tables), explanation requirements, and the cognitive demand of the tasks.

Throughout the study, half of teachers' instructional tasks declined in cognitive demand during implementation (77 out of the 144 total tasks examined in the study). The greatest factor associated with the decline in cognitive demand was that challenges became *nonproblems*, that is, teachers simplified the work required of students. According to Stein et al. (1996), "in many instances, teachers appeared to find it difficult to stand by and watch students struggle, and they would step in prematurely to relieve them of their uncertainty and (sometimes) emotional distress at not being able to make headway" (p. 480). The maintenance of high-level cognitive demand was only 42%

(Stein et al., 1996). This demonstrates that teachers did not implement tasks at the same cognitive levels as they had been set up.

Henningsen and Stein (1997) identified various factors that can attribute to the decline in cognitive demand from task setup to implementation. When examining the decline from *doing mathematics* to *procedures without connections* (high to low cognitive demand tasks, $n = 8$), this study showed that the greatest contributors were that challenges becoming nonproblems (100%), too much or too little time for students to work (75%), and focus shifting to correct answers (75%). Teachers implemented 22 of the tasks at the of *doing mathematics*; among the factors contributing to the implementation of such tasks at a high cognitive level, the three largest were that tasks built on students' prior knowledge (82%), student spent an appropriate amount of time (77%), and sustained pressure for explanation and meaning by teachers (77%). Interestingly, the amount of time spent by students was the second largest contributor in both instances, and emphasis on either correct answers or on mathematical reasoning was the third largest contributor. These findings link the time students spend on a task and the expected responses of teachers (answers vs. explanations) to the decline or maintenance of cognitive demand.

Since the conclusion of the QUASAR project, various researchers have investigated the role of PD in improving mathematics teachers' knowledge of mathematical tasks and levels of cognitive demand. Arbaugh and Brown (2005) sought to engage high school mathematics teachers in "an initial examination of their teaching in a way that is non-threatening and, at the same time, effectively supports the teachers' development of pedagogical content knowledge" (p. 500). The researchers found that engaging high school mathematics teachers in PD that incorporates the levels of cognitive demand (as identified in the TAG) can support teachers in thinking critically about the tasks they give to students and the potential learning that follows. Additionally, participating teachers seemed to change the types of tasks they selected and implemented over the course of the 8-month study; however, no causal relationship could be made linking teachers' task use to the influence of the PD due to the research design (Arbaugh & Brown, 2005).

Boston and Smith (2009) examined high school mathematics teachers participating in the Enhancing Secondary Mathematics Teacher Preparation (ESP) project. Their goal was "to improve the quality of field experiences for preservice mathematics teachers by ensuring that the classrooms of their mentor teachers provided a 'reinforcing culture' of quality mathematics instruction" (Boston & Smith, 2009, p. 128). The ESP project was a series of PD sessions that provided teachers opportunities to engage in mathematical tasks and analyze them based on their cognitive demands. Boston and Smith collected data over a period of 5 days following the PD and found that the cognitive demand of the tasks that teachers selected and implemented increased from fall to winter, showing that high school mathematics teachers can improve the selection and implementation of mathematical tasks after participating in PD focused on the Stein and colleagues (2009) MTF and TAG.

In Nie et al.'s (2013) investigation of middle school mathematics teachers' use of *Standards*-based and traditional textbooks, two of the primary research focuses were on the use of instructional tasks and learning goals. The researchers conducted classroom observations and interviews from teachers of 7 schools implementing Connected Mathematics Program (CMP) textbooks and 7 comparable schools implementing non-CMP texts.

Rather than analyzing the cognitive demand of tasks using the Stein et al. (2009) TAG, Nie et al. (2013) used a different, yet similar, four-category system to rate the cognitive demands of a lesson's learning goals. The system included two low-cognitive demand levels, *remembering* and *practicing*, and two high-cognitive demand levels, *understanding* and *analyzing*. Observers coded both the intended learning goals (what teachers hoped to achieve through the lessons) and implemented learning goals (the goals that observers noticed during each lesson) of each lesson. Comparative results showed that non-CMP teachers planned significantly fewer lessons with high-level learning goals and CMP lessons were implemented at high levels significantly more often than non-CMP lessons. However, the percentage of high-level implemented CMP lessons was significantly fewer than the percentage of high-level planned CMP lessons (Nie et al., 2013). These results support QUASAR research showing that the cognitive demands of tasks (or lessons) tends to decline during implementation, but also show that *Standards*-based curricula may have greater potential for high-level implementation.

Nie et al. (2013) also explored teachers' reasons for choosing instructional tasks by interviewing teachers after they implemented each lesson. Most teachers in the study gave one of two reasons: (a) either teachers followed the guidance of their textbooks or (b) teachers chose tasks to respond to their students' learning needs. Remarkably, the percentage of the former was significantly greater in CMP lessons and the percentage of the latter was significantly greater for non-CMP lessons (Nie et al., 2013). This result, according to Nie et al., was perhaps linked to the high-level planning and implementation of CMP lessons because the *Standards*-based curricula emphasized the development of concepts. Though Nie et al. have already investigated teachers' reasons for selecting tasks, there is still need for the present study for several reasons. First, the majority of teachers in their study provided "brief responses consisting of a single sentence or a few phrases" (p. 705) when asked to explain their task selection, whereas the present study addressed this issue in much greater depth. Second, Nie et al. focused on task selection only, whereas I investigated teachers' reasons for task selection, planning, setup, and implementation. Third, Nie et al. researched middle school teachers, who may have different perspectives than high school teachers, especially those currently involved in PD programs.

In another investigation of textbook curricula, Tran and Tarr (2018) analyzed statistics tasks in 3 mathematics textbook series for Grades 9–12: one traditional series, with probability and statistics in separate chapters or courses (following the Algebra 1, Geometry, Algebra II sequence); one integrated series, with probability and statistics contained within investigations in algebra and geometry; and one hybrid series, with probability and statistics content situated within related algebra and geometry lessons. Tran and Tarr coded 582 tasks within the 3 textbook series based on the *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report* framework (Franklin et al., 2007), including the *statistical investigation components* (formulating questions, collecting data, analyzing data, and interpreting results) and the *statistical developmental levels* (*GAISE* developmental levels of A, B, and C) of each. Tasks were also assigned a level of *mathematical complexity* (low, moderate, or high) based on the amount and type of work required by students to complete each task (similar to the TAG levels of cognitive demand—low complexity may be equivalent to *memorization* or *procedures without connections*, and so on).

Tran and Tarr (2018) found that none of the 582 tasks from the 3 textbook series required students to generate their own statistical questions. Almost every task involved collecting data, however, they did not include opportunities for students to plan data collection procedures to control and reduce variability. Rather, most tasks involved simple experiments or classroom censuses and were coded at low statistical levels (level A in the *GAISE* framework). Analyzing data tasks were scored at higher levels among the hybrid and integrated textbook series (i.e., more B and C scores than A scores), suggesting that the nontraditional texts advanced beyond low-level data collection and analysis. Most tasks in each textbook series were at level B for interpreting results. In terms of mathematical complexity, the distributions for the traditional and hybrid texts were nearly identical; however, the integrated text differed in that fewer tasks were coded as low complexity and more tasks were coded as high complexity. As a whole, the cognitive demands (measured in terms of mathematical complexity) of tasks in the integrated series were higher than those in the other two texts, suggesting that students have greater opportunities to develop statistical concepts and understanding through the integrated text (Tran & Tarr, 2018).

A study by Ni et al. (2018) is unique from the others reported in this chapter; not only did the researchers investigate the cognitive demands of mathematical tasks and their connection to students' cognitive outcomes (e.g., problem solving skills), but they also attempted to link them to students' affective outcomes (e.g., interest in learning mathematics and mathematical dispositions). Ni and colleagues video-recorded classroom instruction in 30 Grade 5 classrooms in a single district in China and coded instructional tasks using an adapted version of the Stein et al. (2009) TAG. Though it is not made clear why in the article, it is surprising that the researchers chose to omit the *doing mathematics* level from the TAG and only include the first three levels of cognitive demand. Tasks were also coded based on the representations available (symbolic, figural, hands-on manipulative) and whether the task called for multiple solution methods.

Using a hierarchical linear model, Ni et al. (2018) found that high cognitive demand tasks were associated with increased interest in learning and classroom participation and enhanced students' views of learning mathematics through inquiry. Multiple representations also predicted students' improvement in solving complex mathematics problems. These findings suggest that students' mathematics affect benefited from the use of high cognitive demand tasks (Ni et al., 2018), contributing to the wealth of literature supporting the use of such tasks. However, neither high cognitive demand tasks nor tasks with multiple solution strategies were positively associated with cognitive learning outcomes in the study. Ni and colleagues claim that the lack of a positive relationship might have been due to the fact that students' cognitive outcomes were measured using a scheme that was not based on the local curriculum. However, the indication of student affective gains is a positive result that contributes to the list of reasons to support high cognitive demand task use.

Two trends in the findings reported above are that (a) tasks tend to decline in cognitive demand during the implementation phase and that (b) PD can enhance teachers' ability to critically analyze tasks and implement them at higher cognitive levels. The QUASAR (Silver & Stein, 1996; Stein et al., 1996) and ESP (Boston & Smith, 2009) projects yielded evidence to support the former, whereas Nie et al. (2013) also found that tasks from *Standards*-based curricula were more difficult to implement at high levels than tasks from alternative materials. Regarding the latter, studies by Arbaugh and Brown (2005) and Boston and Smith (2009) highlighted the importance of PD in shaping the ways that teachers think about and use instructional tasks. Inviting teachers to examine instructional tasks and consider the potential they offer students to learn mathematics can lead to enhanced task selection and implementation (Boston & Smith, 2009).

Beliefs and Knowledge as Factors Influencing the Use of Mathematical Tasks

Teacher education literature applies numerous definitions to *knowledge* and *beliefs* and develops the relationship between it and knowledge in various ways, though I consider knowledge and beliefs as two separate (but not mutually exclusive) entities. Many scholars have studied and identified connections between knowledge and beliefs, some writing that one influences the other (e.g., Lee et al., 2019), that one is a subset of the other (Philipp, 2007), or that they are "indistinguishable" (Beswick, 2011, p. 128). Regardless of the way in which knowledge and beliefs are defined or conceptualized, there is general agreement that both are influential in shaping teachers' instructional practices. It is in this way that the literature related to teacher knowledge and beliefs has implications for mathematics teachers' selection, planning, and implementation of mathematical tasks.

My interpretation of beliefs aligns with that of Ambrose (2004), Beswick (2006), and Green (1971). Ambrose considered beliefs as having four components: (a) their origins, (b) their influence on how experiences are interpreted, (c) their individual connections forming belief systems, and (d) how they change. In this framework, beliefs originate from two sources: "emotion-packed experiences and cultural transmission" (Ambrose, 2004, p. 93). Emotional experiences shaping teachers' beliefs about mathematics teaching and learning might occur during their K–12 schooling; for example, learning multiplication tables influences some preservice teachers to believe that they cannot learn mathematics (Ambrose, 2004). The way in which teaching is relatively similar within each culture (Stigler & Hiebert, 1999) and past experience might also influence preservice teachers' beliefs that mathematics should be taught as the memorization of facts and application of routinized procedures.

Ambrose (2004) explained that beliefs have a "filtering effect" (p. 94), an influence on a person's interpretation of new experiences. This means that a person's existing beliefs influence their perception of new experiences and how they might interact within these experiences. Ambrose gave an example of a colleague whose preservice teachers worked with kindergarten students: the experience was meant to show preservice teachers at the start of the semester the wealth of prior, informal knowledge the kindergarten students brought to school. However, the experience left the preservice teachers astounded by how much the kindergarten teacher taught students in a single week. However, their beliefs that mathematics knowledge comes from formal schooling potentially influenced this reaction. The preservice teachers attributed the teaching to the impressive amount of knowledge developed by the kindergarten students at the end of the week rather than the knowledge they already possessed. The filtering effect described by Ambrose helps to explain the differences in teachers' responses to teacher training (Ambrose, 2004) and the introduction of new mathematics curricula (Philipp, 2007); teachers' preexisting beliefs about the teaching and learning of mathematics might hinder their ability to modify their beliefs through new experiences.

Multiple authors agree that beliefs are interconnected to each other, forming belief systems (Rokeach, 1968) or clusters (Green, 1971). Ambrose drew from Green's notion that, though individual beliefs are connected to form belief systems, these systems sometimes exist in isolation of each other; that is, two belief systems may lead to seemingly contradicting viewpoints. For example, a lack of creative experiences with mathematics may cause a disconnect between preservice teachers' belief systems valuing creative learning and their belief systems concerning mathematics (Ambrose, 2004). The last of Ambrose's components, changes in belief systems, occur when people (a) encounter emotionally powerful experiences, (b) immerse themselves in communities through the process of cultural transmission, (c) reflect on their beliefs and reveal those that might be hidden, (d) encounter experiences that work to connect isolated belief systems, and (e) when a belief reversal occurs.

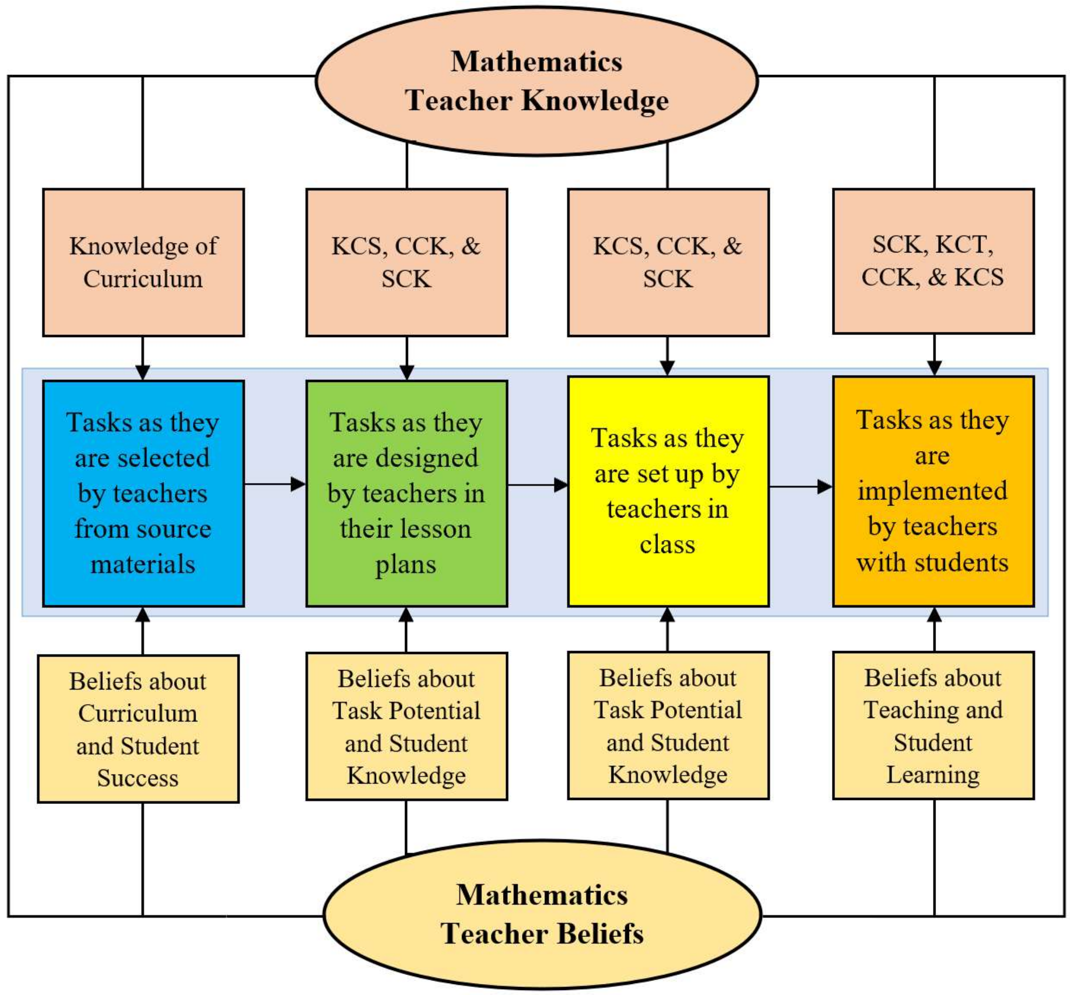
Beswick (2006) also drew upon Green's (1971) description of beliefs and belief systems, suggesting that beliefs are strengthened by their connections to other beliefs and that beliefs are formed on some basis and exist within a context. Beswick described two bases on which beliefs are held: evidence or non-evidence. A belief formed on the basis of evidence might be a belief supporting a particular pedagogical approach after students' test scores increase. A belief formed on the basis of non-evidence might be a preservice teacher's belief formed through interactions with his or her mentor teacher; in this case, the belief may be held because of emphasis from an authority figure. Beliefs formed on the basis of evidence are susceptible change with the introduction of evidence to support the change. However, beliefs formed on non-evidence are unlikely to be changed by new evidence. Context influences the relative centrality of beliefs and supports the notion of contradictory beliefs among unconnected belief systems; one of two unconnected systems may be more central than the other in a particular context. For example, a teacher might communicate a belief in teaching mathematics using manipulatives in exploratory ways during a PD workshop but might also strive to maintain control and limit students' ability to explore in the context of the classroom (Beswick, 2006).

Teachers' Knowledge and Beliefs Influence Their Use of Tasks

The following review of the literature pertaining to mathematics teachers' knowledge and beliefs cites evidence that both influence teachers' selection of tasks from source materials, task planning, task setup, and task implementation with students. These relationships are depicted in Figure 2.

Figure 2

The Influence of Teacher Knowledge and Beliefs on Mathematics Teachers' Task Use



Note. Three components of Mathematics Teacher Knowledge, influence tasks as they are designed by teachers in their lesson plans (Phase 2) and tasks as they are set up by teachers in class (Phase 3), and tasks as they are implemented by teachers with students (Phase 4): Knowledge of Content and Students (KCS), Common Content Knowledge (CCK), and Specialized Content Knowledge (SCK). Knowledge of Content and Teaching (KCT) also influences tasks as they are implemented by teachers with students (Phase 4).

At the first phase (tasks selected by teachers from source materials), knowledge of curriculum (Hill, Blunk, et al., 2008) and beliefs about which curricular materials effectively help students learn mathematics (Carson, 2011; Collopy, 2003; Philipp, 2007) influenced the curricula and tasks selected by teachers. This phase also incorporates beliefs about what it means for students to learn and be successful in school mathematics (Collopy, 2003). At the second and third phases (tasks as planned and set up by teachers), Knowledge of Content and Students (KCS) supports teachers in anticipating how students will engage with tasks (Ball et al., 2008), whereas Common Content Knowledge (CCK) and Specialized Content Knowledge (SCK) influence teachers' planning (Charalambous, 2010). Lee (2019) found evidence that the combination of teachers' beliefs about students' prior knowledge and beliefs about task potential shaped the ways in which teachers not only selected tasks but also the planning and setup that followed.

A variety of knowledge influences tasks at the fourth phase (tasks as implemented in the classroom). Charalambous (2010) found evidence linking both CCK and KCS to the cognitive demands of enacted tasks, and Chapman's construct of Mathematical Task Knowledge for Teaching (MTKT) includes elements CCK and KCS, specifically addressing teachers' knowledge of tasks throughout the stages of the MTF. Beliefs regarding teaching as telling (Ambrose, 2004; Philipp, 2007; Romagnano, 1994), beliefs about how mathematics should be taught (Beswick, 2011; Philipp, 2007; Ross et al., 2002), and beliefs about mathematics as a discipline (Beswick, 2011) influenced the instructional decisions made by teachers during instruction.

The knowledge and beliefs described in the previous paragraphs influence teachers' selection, planning, and enactment of mathematical tasks. Therefore, each of these findings was incorporated into Figure 1 on page 50 to produce an adjusted version that captures relevant teacher knowledge and beliefs at each phase. Figure 2 depicts these influences on the stages of teachers' task use with arrows to indicate which stage is affected by a knowledge type or belief. Knowledge and beliefs are also linked in the sense that knowledge is a stronger form of belief in what is thought to be true (Clement, 1999; Philipp, 2007; Wilson & Cooney, 2002); this relationship is also depicted in Figure 2. The following subsections describe the body of literature forming the basis for this framework.

Mathematics Teacher Knowledge

Mathematics teaching requires a broad scope of knowledge, including knowledge of mathematics content, pedagogy, and knowledge of students and their learning (Monk, 1994; Ball et al., 2008). Before discussing implications of mathematics teacher knowledge on the MTF, it is necessary to describe Shulman's (1986) Pedagogical Content Knowledge (PCK) and Ball et al.'s (2008) Mathematics Knowledge for Teaching (MKT). Theoretical arguments can be made for how PCK and MKT influence tasks at the four task phases. However, empirical results from various studies help to support these claims.

Shulman and Pedagogical Content Knowledge. Knowledge of subject matter content alone is not sufficient for effective teaching; effective teachers draw on various forms of knowledge (Monk, 1994; Stein et al., 1996). Content knowledge, though not encompassing the entire range of knowledge needed for teachers to successfully implement cognitively demanding tasks with students, does comprise of part of this knowledge base. In an examination of literature on teacher knowledge, Shulman (1986) writes that literature at that time did not answer questions about how teachers make pedagogical decisions and where that knowledge comes from. Though research on teaching and learning at the time addressed issues of student learning, it failed to address issues of teacher learning.

To address gaps in the literature of the day, Shulman (1986) suggested three types of content knowledge: subject matter content knowledge, pedagogical content knowledge (PCK), and curricular knowledge. Subject matter content knowledge (e.g., knowledge mathematics) does not simply refer to knowledge of facts and constructs. Teachers should possess comprehensive and rich knowledge of subject matter content and how it extends to various disciplines and applications. From a mathematical task standpoint, this implies that teachers should understand and be able to demonstrate the level of academic rigor expected of their students. They should understand the mathematical connections concerning the content they teach and be able to communicate mathematical ideas, concepts, and their connections to support students' engagement with tasks. This is evidenced by Stein et al. (2007) in their review of curriculum research, describing a study in which elementary teachers did not enact mathematical tasks at high levels because they failed to appreciate the mathematical complexity underlying activities suggested in curricular materials. In such cases, teachers' lack of subject matter content knowledge can prevent tasks from being enacted in ways that challenge students to engage in higher order thinking.

Pedagogical content knowledge, PCK, is “the ways of representing and formulating the subject that make it comprehensible to others” (Shulman, 1986, p. 9). This also includes knowledge of how students learn particular concepts, the relevant prior knowledge students bring with them and how it relates to the content, and how to structure learning opportunities for students in ways that will make the content accessible. This is certainly related to research on mathematical tasks, indicating that prior experience and teachers’ questioning practices influence a task’s level of implementation (Henningsen & Stein, 1997; Stein et al., 2007). Pedagogical content knowledge varies from subject matter content knowledge in the sense that PCK contains information that is unexpected and even unnecessary for others within a subject to grasp (Shulman, 1986). For example, a high school mathematics teacher must know common student misconceptions in high school algebra and how to handle them appropriately. This type of knowledge is not needed in the work of a mathematics researcher studying algebraic topology.

Shulman’s (1986) construct of curricular knowledge is the knowledge of the instructional materials available to teach content and the knowledge of how to select the appropriate materials to accomplish various goals. Shulman dissected curricular knowledge into two parts: lateral curriculum knowledge, the knowledge of how grade-level content relates to content in other subjects, and vertical curriculum knowledge, the knowledge of how grade-level subject matter content relates to what students already know and how it relates to they will learn in the future. Curricular knowledge is relevant to the work of Stein et al. (1996) in task enactment in the sense that part of mathematics curriculum involves the tasks with which students engage. Mathematics teachers require the curricular knowledge to be able to assess the quality of an instructional task and determine whether a task is appropriate for their students.

Mathematical Knowledge for Teaching. Hill, Ball, and Schillings (2008) described the construct Mathematical Knowledge for Teaching (MKT), an extension of Shulman's (1978) PCK. The MKT framework includes PCK and divides it into three sections:

- Knowledge of Content and Students (KCS), content knowledge and the "knowledge of how students think about, know, or learn this particular content" (p. 375),
- Knowledge of Content and Teaching (KCT), knowledge of how to engage students in learning through teaching moves, and
- Knowledge of curriculum, as defined previously by Shulman.

The definition of KCS, according to Ball et al. (2008), is "knowledge that combines knowing about students and knowing about mathematics" (p. 401). This includes knowledge of students' thinking, knowledge of what students will find difficult and confusing, and knowledge of students' conceptions and misconceptions. Such knowledge guides the decisions and actions that teachers make regarding students, including which tasks to select and the anticipation of how students might interact with it. Knowledge in the KCS domain is at the intersection of mathematics content and students (Ball et al., 2008). Similarly, KCT "combines knowing about teaching and knowing about mathematics" (p. 401), representing the intersection of mathematics content and teaching. The domain KCT contains the knowledge that mathematics teachers use to inform pedagogical decisions, such as which mathematical representations to use when teaching a particular topic. Orchestrating effective mathematics discussions requires strong KCT, as teachers must decide which student responses to follow up on and which to return to later. Curricular knowledge, referred to as "knowledge of content and curriculum" (Ball et al., p. 403), retains a similar meaning to Shulman's (1986); however, Ball and colleagues include it within PCK rather than as a separate form of knowledge. The way in which Ball et al. conceptualize PCK distinguishes between knowledge for teaching, knowledge for learning, and knowledge of content and the various intersections between them.

The second part of the MKT framework, Subject Matter Knowledge, is comprised of three sections:

- Common Content Knowledge (CCK), an interpretation of what Shulman meant by subject matter knowledge,
- Specialized Content Knowledge (SCK), involving the content knowledge related to teaching tasks, and
- knowledge at the mathematical horizon (Hill, Ball, & Schillings, 2008).

Ball et al.'s (2008) notion of CCK follows Shulman (1986) in that CCK is the knowledge of how to use and apply mathematics in daily, real-world settings and knowledge of mathematical objects and constructs that might be used by mathematicians. Though CCK refers specifically to knowledge of mathematics outside the realm of teaching, this knowledge is needed by teachers; they must understand both the procedures and concepts of the mathematics for themselves in order to teach this content to others effectively. Moreover, mathematics teachers must possess the knowledge necessary to identify students' mathematical errors and the knowledge to support the correct use of mathematical language (Ball et al., 2008). Ball et al. defined SCK as "the mathematical knowledge and skill unique to teaching" (p. 400); this includes how to handle student errors and misconceptions and how to determine whether students' nonstandard algorithms or approaches apply in a given context and whether such approaches apply more broadly. This type of knowledge is applied in the daily work of teaching and is unnecessary in other professions (Ball et al., 2008). The third domain of Subject Matter Knowledge, knowledge at the mathematical horizon, is the knowledge of how mathematics content is related to other content in the curriculum. This is similar to Shulman's (1986) vertical curriculum knowledge; for example, this includes the knowledge of how first grade mathematics content related to what will be studied in third grade and informs teachers to make decisions that will support students' future mathematics learning (Ball et al., 2008).

Knowledge and Task Selection. The connection between curricular knowledge and teachers' task selection is evident in a study by Hill, Blunk, et al. (2008), who investigated MKT and Mathematical Quality of Instruction (MQI), "a composite of several dimensions that characterize the rigor and richness of the mathematics of the lesson, including the presence or absence of mathematical errors, mathematical explanation and justification, mathematical representation, and related observables" (p. 431). Lauren, one of the mathematics teachers who demonstrated strong MKT and high MQI in the study, highlights her skill in task selection and sequencing that are affordances of her MKT. According to Hill, Blunk, et al., Lauren's ability to select and sequence tasks is supported by her knowledge of mathematics teaching, benefitting her use of tasks as written in curricular materials and tasks as planned. Moreover, her strong MKT helped her to draw on a variety of curricular materials to piece together tasks that develop mathematical ideas in a coherent way. Within the MKT framework, it appears as if Lauren's knowledge of curriculum (or curricular knowledge) supported her effective use of written tasks. This knowledge is evident in her selection of high cognitive demand curricular materials (e.g., tasks adapted from *Investigations in Number, Data, and Space).* Lauren's teaching also benefited from her strong MKT, as her lessons included rich mathematics and language (Hill, Blunk, et al., 2008). In contrast, some teachers who scored lower on the MKT assessment did not select tasks from vetted, standards-based curricula.

Knowledge, Task Planning, and Task Setup. There is evidence that teachers' task planning, and consequently, task setup and implementation, is limited to the extent of their mathematics knowledge. That is, teachers are unlikely to plan or set up a task that explores mathematical knowledge beyond what they possess. Charalambous (2010) conducted research on the influence of mathematics teachers' knowledge in selecting, planning, and implementing tasks, proposing that "there exists a relationship between teachers' mathematical knowledge for teaching (MKT) and their decisions and actions during the three phases of task unfolding" (p. 248). This premise is support theoretically in the sense that MKT includes knowledge related to lesson planning, determining the quality of tasks and modifying them for instruction, successfully teaching both concepts and procedures to students, and addressing the mathematical work and reasoning done by students (Charalambous, 2010). Empirical evidence for the linking of MKT to the MTF includes the work of Ball and colleagues (2008) and a study by Hill et al. (2005), providing evidence suggesting that teachers with high MKT supported students in greater learning gains than teachers with average MKT.

Charalambous (2010) focused on case studies of two elementary teachers with differing MKT and included data from 18 lessons (9 for each teacher), analyzed from the perspective of the MTF and with the use of the TAG by Stein and colleagues (2009). Additionally, both teachers completed a written Learning Mathematics for Teaching (LMT) test, measuring CCK and SCK of numbers and operations, geometry, and algebra; the two teachers, Karen and Lisa, scored in the 93rd and 35th percentile, respectively (Charalambous, 2010). Teachers participated in task-based interviews, interviews during which they solved mathematical tasks, and post-lesson interviews to elicit information regarding teachers' mathematics background, experiences, and beliefs. Charalambous (2010) reports that "Lisa set up most of the tasks (i.e., about 83%) at a low cognitive level by mainly asking students to recall and apply rules and algorithms" (pp. 257–258), implying that this finding translates between the phases of the MTF (i.e., teachers with lower MKT scores are less likely to select, plan, set up, *and* implement tasks at high cognitive levels). Charalambous connected these findings to the task-based interviews taken by each teacher, highlighting the procedural approach Lisa took to solve mathematical problems as opposed to the conceptual, representational approach taken by Karen (e.g., Lisa translating problems into sequential statements such as “1/5 of 1/2,” using the word “of” as a trigger word meaning multiplication, a procedural technique not grounded in mathematical concepts, see p. 268).

Knowledge and Task Implementation. Hill and Ball (2004) suggested that the nature of mathematics teachers’ knowledge might be more influential than how much knowledge they hold, that is, “whether teachers' knowledge is procedural or conceptual, whether it is connected to big ideas or isolated into small bits, or whether it is compressed or conceptually unpacked (Ball & Bass, 2000; Ball, Lubienski, & Mewborn, 2001; Ma, 1999)” (p. 332). This speaks to the MTF in the sense that teachers with a limited, procedural knowledge of mathematics might find it difficult to implement tasks at high cognitive demand and teachers with strong conceptual understanding of the mathematics they teach may be more likely to support students in high-level task enactment. For example, Charalambous’ (2010) investigation of two elementary school mathematics teachers, Lisa and Karen, found that the nature of each teacher’s mathematical knowledge was similar to the ways in which they enacted tasks in each of their classrooms. When solving a problem that involved finding the fractional part of a fraction, Karen used a representational approach by drawing a rectangle and partitioning a fraction of it into equal parts. Lisa, however, called upon her knowledge that the word “of” meant multiplication and used a procedural approach to solve the problem. Moreover, Charalambous found that Karen's approach to reason through problems and Lisa's procedural approach were evident in their teaching.

Charalambous (2010) hypothesized that MKT supports teachers' use of mathematical representations to enhance students' understanding of the concepts underlying procedures. The knowledge that teachers possess appears to influence both the ways in which teachers solve mathematical problems but also the ways in which they interact with their students. Moreover, teachers' MKT enhances their ability to not only provide explanations for mathematical concepts underlying procedures but also their ability to scaffold students' development of such explanations. This is perhaps because teachers' ability to communicate mathematically is based on their knowledge of mathematics, teaching, and the combination of the two. Reasonably, it is unlikely that students will develop deeper conceptual understandings of mathematical procedures than those understandings possessed by their teachers; however, some exceptions might arise when students do additional work and seek additional resources beyond their teachers' instruction. Charalambous also provided evidence that MKT influences teachers' ability to facilitate students' sense making of mathematics through their responses to what students say and do in class. Therefore, there is evidence that strong MKT is apparent in the work of teachers who engage and support their students in the enactment of high cognitive demand tasks; the potential for student learning is not only based on the task itself at each phase of the MTF, but also on the knowledge of the teacher.

Chapman (2013) echoed the voices of Charalambous (2010) and others, providing further insights to the role of teachers' knowledge in their instructional practices, specifically in the enactment of high cognitive demand tasks. Chapman writes that

> it is the teacher and students who give [tasks] life based on how they are interpreted and enacted in the classroom. The teacher is critical in shaping the lived task and directing students' activities so that students have opportunities to engage meaningfully in mathematics through them. (p. 1)

The ways in which a teacher interacts with his or her students, including interactions that shape the nature of a task (e.g., whether a task is open-ended) or change its level of cognitive demand, Chapman attributes to teachers' content knowledge, knowledge of students, goals, teaching style, and beliefs. For example, interactions between teachers and students that influence task implementation are affected by teachers' KCS (Ball et al., 2008). Deeper knowledge of how students interact with mathematics content helps mathematics teachers to make appropriate instructional moves that support students' task engagement without lowering the cognitive demand. Additionally, teachers with stronger CCK can guide their students toward mathematically precise understanding throughout task engagement, linking Chapman's ideas to the work of Ball et al.

A construct that Chapman (2013) referred to as "mathematical-task knowledge for teaching" (p. 1) includes the knowledge required of teachers to select and modify tasks to promote conceptual understanding and enhance their potential to influence student learning. According to Chapman, the knowledge comprising of mathematical-task knowledge for teaching (hereafter referred to as MTKT) consists of six components:

1. knowledge of task characteristics that promote students' mathematical reasoning and conceptual understanding,
2. the ability to detect and create tasks that are mathematically rigorous and worthwhile for their students,
3. knowledge of the TAG and levels of cognitive demand (Stein et al., 2009) and how those relate to the goals for the task,
4. an understanding of students' prior knowledge, experiences, and interests,
5. an understanding of the relationship between selected, planned, and implemented tasks and their influence on students' ability to learn mathematics, and
6. knowledge of how to enact high cognitive demand tasks without taking over the work and thinking of students.

Each point on Chapman's list is arguably necessary to select, plan, and implement mathematical tasks at high cognitive levels.

Each of the six components of MTKT influence the MTF at various stages. For example, components 1–2 are influential in task selection or creation; the knowledge of tasks that support students' mathematical reasoning and understanding, for instance, may influence teachers to choose some tasks for implementation and others not. However, teachers may select tasks for a variety of reasons, including surface-level features such as whether a task involves graphing (Arbaugh & Brown, 2005; Stein et al., 1996). Though not the only factor, the aforementioned knowledge may help to distinguish teachers who select high cognitive demand tasks frequently from those who might not. Moreover, knowledge of goals for the task and students' prior knowledge (components 3–4) may help teachers in the task planning, each of which might inform teachers' decisions when making task modifications. Teachers' planning of tasks should aim to meet the goals for the task and goals for student learning, whereas an understanding of students' prior knowledge should work to inform such goals and how the teacher will work to help students meet them. Component 5 encompasses all three task phases in but specifically focuses on how they work together to influence student learning, the outcome of implemented tasks in the MTF. As evidenced in the study by Henningsen and Stein (1997), teachers' knowledge and ability to enact tasks without removing the thinking done by students is also crucial in task enactment (component 6 of MTKT proposed by Chapman).

Mathematics Teacher Beliefs

Just as mathematics teachers' knowledge is influential to their practice, beliefs play a similar role in shaping the ways in which they interact with mathematics content, students, and tasks. For example, some preservice teachers spend the majority of their lives in the K–12 school setting and believe that this experience sufficiently prepares them for what to do and expect as teachers, experiencing the "optimistic bias" (Ambrose, 2004) that their experience in school settings and with children alone will guide them to be successful teachers. The previous section of this literature review alone makes evident the plethora of knowledge required to teach for student understanding; some preservice teachers underestimate the importance of such knowledge or might be unaware of it entirely (Ambrose, 2004). Beliefs may potentially shape teachers' decisions about what tasks to use (Collopy, 2003; Philipp, 2007), how tasks should be modified based on students' prior knowledge (Lee et al., 2019), and how tasks should be enacted with students (Philipp, 2007; Romagnano, 1994). Moreover, teachers' beliefs about teaching and enacting mathematical tasks may also be related to their beliefs about mathematics as a discipline (Beswick, 2011). Like the section on mathematics teacher knowledge, the review of literature concerning beliefs is organized into three sections corresponding to the phases of the MTF: (a) task selection, (b) task planning, and (c) task enactment, each of which is described in the following paragraphs.

Beliefs and Task Selection. Philipp's (2007) review of literature on teachers' beliefs included a study involving mathematics teachers' beliefs and their inclinations toward mathematics curriculum, highlighting the impact of their beliefs on their selection and implementation of tasks. Collopy (2003) revealed the way in which teacher beliefs influence the first and third phases of the MTF (tasks as they appear in curricular materials and tasks as enacted in the classroom). Collopy's study involved two elementary school teachers' changes in beliefs through their first year implementing *Investigations in Number, Data, and Space*, a reform-oriented curriculum. Though the two teachers came from similar backgrounds and shared similar experiences, their identified beliefs and implementation of the new curriculum varied considerably due to their beliefs about effective curricula and beliefs about teaching and learning mathematics (Philipp, 2007).

The case of Ms. Clark in the Collopy (2003) study presented a teacher with the traditional belief that "computational speed and accuracy distinguished successful students from unsuccessful students, that her students needed to learn the rules of mathematics, and that mathematics topics should be presented systematically from easier to harder" (Philipp, 2007, p. 288). She also held the belief that understanding meant "the memorization and correct execution of standard algorithms" (p. 288). These beliefs led her to modify the reform curriculum, which emphasized other skills and knowledge, by removing opportunities for students to actively problem-solve and discuss mathematics with others. Persistent frustration eventually led her to abandon the curriculum entirely; throughout her time with the text, she reportedly spent 2.8% of observed time teaching for conceptual understanding (Philipp, 2007). However, in the case of Ms. Ross, the adaptation of the new curriculum influenced a shift in pedagogy and instruction from mostly procedures to emphasis on conceptual understanding (Philipp, 2007). Collopy found that Ms. Ross believed in supporting students' mathematics confidence and building on their prior knowledge; Ms. Ross implemented the new curriculum carefully and methodically, learning more about instruction and about mathematics throughout the process. These two cases provide evidence that teachers' beliefs about mathematics and mathematics teaching influence their beliefs toward mathematical tasks as represented in curriculum materials, their decisions about which curricula should be implemented, and how effectively curricula are implemented (Philipp, 2007).

Another example of beliefs' influence on the selection of mathematical tasks is the case of Anna, a third example from the MKT and MQI study by Hill, Blunk, et al. (2008). Though Anna's MKT was average, two factors negatively influenced the mathematical quality of her instruction: her beliefs about the importance of mathematics being fun and her persistent use supplemental mathematical activities rather than using lessons from her textbook, which resulted in her students spending time engaged in nonmathematical activity (a cognitive demand level below *memorization* according to the TAG). Hill, Blunk, et al. postulate that her limited content knowledge might have influenced the lack of mathematical rigor in her lessons and her beliefs about mathematics being fun might have led to the selection of a task that focused on primarily nonmathematical activity.

Along the same lines as Hill, Blunk, et al. (2008), Carson (2011) additionally acknowledged the apparent relationship that exists between beliefs and the selection of mathematical tasks. Carson studied high school mathematics teachers' beliefs about Exploratory Learning Activities (ELAs): this study investigated 5 high school mathematics teachers' changes in beliefs after participation in a one-day workshop on ELAs. Interviews with participants yielded data suggesting that they viewed ELAs as student-centered, having the potential to increase students' understanding and engagement, and being able to provide active, hands-on experience; however, participants were unable to translate their ELA views into their practice (Carson, 2011). Carson attributed this to teachers' "restrictive beliefs" (Carson, 2011, p. 149), beliefs that they could only enact ELAs occasionally in their practice and not as a dominant pedagogy. Restrictive beliefs also included teachers' beliefs that ELAs were sophisticated activities that required more planning time than they had available (although Chapman argues that the ELAs were not the large projects that the teachers perceived them to be). Carson's study suggests that teachers' restrictive beliefs about the nature of ELAs inhibited their ability to change their practice through the activities (tasks) that they selected.

Beliefs, Task Planning, and Task Setup. Lee et al. (2019) investigated the influence of teachers' beliefs of prior knowledge, their beliefs in task potential to enhance students' knowledge, and their beliefs about how students' prior knowledge and task potential influence their instructional practice. The study included data from 54 teachers participating in the first year of a 3-yr PD program focusing on Algebra I teachers' knowledge and implementation of ideas from the *Common Core State Standards for Mathematics* (National Governors Association Center for Best Practice & Council of Chief State School Officers, 2010). Data included teachers' engagement in solving tasks, lesson planning meetings, classroom observations, and a workshop during which teachers discussed a learning trajectory related to the lesson they taught (Lee et al., 2019). Based on teachers' various beliefs about students' prior knowledge and beliefs about task potential and combinations of the two, Lee et al. identified three belief patterns, each associated with differing teaching characteristics: (a) prior knowledge as a foundation to develop and tasks having the potential to create new knowledge, (b) prior knowledge as an entity and tasks with no potential for new knowledge, and (c) prior knowledge as a foundation to develop but constrained by the prior context and tasks with on potential for new knowledge.

The first of these belief patterns emerged through their use of students' prior knowledge as a way to "open up something problematic for students" (Lee et al., 2019 p. 142). Problematic, in this sense, meant that teachers tended to use students' prior knowledge as an entry point for a new problem situation. Such teachers planned the instructional use of tasks around what students already knew and enacted tasks to create new knowledge from it. The second pattern of beliefs, prior knowledge as an entity and no potential of a task for new knowledge, limited the planning and implementation of tasks; teachers with this belief pattern shaped their tasks around reviewing concepts that they already learned based on the belief that tasks could not develop new knowledge. For example, Mr. Devlon planned and implemented a task a way such that only focused on review of past knowledge; when a student expressed a new concept, Mr. Devlon did not raise the issue with the whole class to promote further knowledge development. Lee et al. concluded that the purpose of using the task was to find answers rather than allow students to develop new knowledge. The purpose of the task and its potential, beliefs of Mr. Devlon, appeared to influence task setup and implementation. The third belief pattern, prior knowledge as an entity and no potential of a task for new knowledge, incorporated a similar view of tasks as limited to promoting factual recall and limited teachers' planning and implementation (Lee et al., 2019).

Beliefs and Task Implementation. According to Beswick (2006), meaningful and lasting change can only occur when teachers' belief systems align with the changes they are trying to make. Teachers must be convinced that the change they are trying to make is valuable and worthwhile. Teachers' instructional practices are closely tied to their beliefs and are difficult to change without first changing their beliefs. Beswick (2006) examined the connections between a sample of 25 high school mathematics teachers' beliefs and classroom environments, including data from surveys of both teachers and students, teacher interviews, and classroom observations. One of the findings of this study was that, despite the various methods of implementation, teachers' beliefs affected their classroom environments and instruction in ways that were noticeable to students. This result suggests that participating teachers' beliefs were associated with the instructional decisions they made, influencing the enacted curriculum as identified by Stein and colleagues (2009) in the third task phase of the MTF (tasks at enacted). Beswick's asserts that teachers' beliefs influence their teaching, a statement that is "widely acknowledged" (p. 17) in the field. That is, teachers attempting to enact mathematical tasks with high cognitive demands are likely to be unsuccessful if their beliefs are consistent with what they are trying to implement.

Ambrose (2004) attempted to understand and change the beliefs of preservice teachers at the onset of their study and training through direct experience working with children in elementary school while simultaneously taking their first mathematics-for-teachers course. Ambrose included a review of literature pertaining to teachers' beliefs, including the notion of a belief system, for example, the connected beliefs that students should have opportunities to be creative that is connected to beliefs about subject matter such as writing or the arts. Some teachers, though they hold such a belief system, may not feel the same way about mathematics because of how they experienced it as children (Ambrose, 2004). Moreover, "beliefs about mathematics teaching and learning are part of a larger system of beliefs that includes beliefs about teaching, generally" (p. 96). Additional preservice teachers' beliefs about teaching identified by Ambrose are that teaching involves presenting content to be memorized (Richardson, 1996) and that teaching is straightforward (Fieman-Nemser et al., 1988). Some believe that teaching is the simple transfer of information from teachers to students (Wideen et al., 1998).

Ambrose's (2004) study included data from 15 preservice teachers involved in taking an experimental mathematics-for-teachers course coupled with their first mathematics course. Data obtained through surveys, interviews, written work, and observational field notes revealed that, throughout their time interacting with students and learning new mathematics, preservice teachers' beliefs about teaching and learning changed; they began to see that teaching is more complicated than simply passing information on to their students (Ambrose, 2004). This general finding applied to the majority of the students involved in the study. The researcher attributed shifts in beliefs about the nature of teaching and learning, student discovery, and students' use of multiple solution strategies to solve problems to the experiences preservice teachers had with children during their courses. Importantly, their experiences "helped them to realize that their understanding of mathematical concepts was essential to their success as teachers" (p. 114), suggesting that beliefs about the knowledge required for teaching are coupled with the knowledge for teaching itself. However, more than half of the preservice teachers involved in the study continued to cling to their beliefs about teaching as telling in practice despite what they had said during interviews, presenting of standard algorithms for the addition of fractions (Ambrose, 2004).

Similarly, in an early effort to positively influence preservice teachers' beliefs regarding teaching mathematics through the use of technology, Zelkowski (2009) designed and implemented an introductory mathematics education course and gauged its influence on preservice teachers' beliefs about teaching and technology. The premise for the course was that the majority of preservice teachers in the area would be placed into traditional teacher-centered classrooms, with procedural tasks and limited use of technology to support their adoption of ambitious teaching practices in this context. Zelkowski designed the introductory teaching with technology course to provide mathematically rigorous learning opportunities for preservice teachers and to support preservice teachers' knowledge and application of the NCTM technology and equity principles.

Three emerging findings appeared from preservice teachers' experiences engaging in course activities: (a) some preservice teachers were challenged to reflect on their beliefs about mathematics teaching, realizing that mathematics can be engaging rather than simply procedure-oriented, (b) preservice teachers' were led to reevaluate their mathematics schooling, revealing that their experiences lacked opportunities to develop deep mathematical understandings, and (c) preservice teachers' beliefs about graphing calculators as purely computational, answer-getting, and graphing tools were changed to incorporate new, expanded beliefs (Zelkowski, 2009). Regarding the first finding, Zelkowski asserted that "the technology course begins the process of challenging and re-molding beliefs about teaching mathematics as involving mostly rote algorithmic skills and memorization" (pp. 78–79); moreover, these changes in beliefs are also influencing changes in practice: preservice teachers reportedly provided evidence that they are reassessing their beliefs about teaching mathematics. These preliminary findings suggest that mathematics teachers' beliefs are influenced and can be changed through experiences engaging in mathematical activity and learning (that is, beliefs are influenced by the enacted curriculum, the third task stage in the MTF).

Another implication for teachers' practice, and therefore, the enactment of mathematical tasks, was evident in a study by Beswick (2011) comparing teachers' beliefs about mathematics as a discipline to their beliefs about teaching mathematics. Beswick investigated whether teachers held differing views of mathematics as a discipline and school mathematics and implications that it might have for teachers' practice. The interest for investigating differences in such beliefs is that the typical work done by students in mathematics classrooms is strikingly different than the work of mathematicians; traditional mathematics classrooms consist of problems selected by teachers, solvable within minutes, and require little cognitive work, whereas mathematicians work with autonomy, spend a considerable amount of time solving a single problem, and often struggle while doing so (Beswick, 2011). Knoll et al. (2004) argued that school mathematics should strive to align more closely with mathematics as done by professionals by including activities such as searching for solution strategies, identifying examples and counterexamples, and the use of argumentation and proof (as cited in Beswick, 2011). Though this appears to be a general belief held by mathematics educators and professional mathematics organizations (e.g., NCTM), typical school mathematics instruction does not follow this approach (Stein et al., 1996; 2009).

Beswick (2011) collected data pertaining to teachers' beliefs using a survey containing items related to beliefs about mathematics and mathematics teaching and learning, semistructured interviews, and classroom observations. Sally, an experienced mathematics teacher with a strong mathematics background, that she taught using a problem-solving approach based on her beliefs about school mathematics rather than her beliefs about mathematics as a discipline. Sally based her beliefs on advancements in the teaching profession that were made throughout her career (Beswick, 2011). Jennifer, a teacher whose beliefs about mathematics teaching and learning and beliefs about mathematics followed a Platonist view (that is, the view that mathematics is an unchanging set of knowledge that is yet to be discovered by students) was less successful in implementing a problem-solving approach than Sally. Beswick concluded that:

> rather than teachers' beliefs about the nature of mathematics as a discipline necessarily influencing their teaching in theoretically consistent ways it appears that, where there is a difference, these beliefs interact with those they hold about school mathematics to influence their beliefs about teaching and learning. (p. 145)

These beliefs, in turn, influence the ways in which teachers implement their instruction, as evidenced by Sally and Jennifer.

Ross et al. (2002) also found evidence that the largest barrier preventing implementation of innovative teaching practices was teachers' beliefs about mathematics teaching. This implies that PD initiatives focusing on the work of Stein and colleagues (2007) and the MTF are unlikely to inspire teachers to incorporate high level tasks unless the PD also addresses and works to change teachers' beliefs as well. Raymond (1997) found that teachers' beliefs about mathematics were more closely related to practice than beliefs about teaching and learning; Raymond identified a teacher, Joanna, whose self-reported mathematics beliefs were traditional whereas her self-reported beliefs about mathematics teaching and learning were nontraditional. Because her teaching practices were traditional in nature, including students sitting at their desks working quietly by themselves, teacher-directed instruction, and strict discipline she "seemed to view her mathematics-teaching practice in terms of what she wanted to do, or thought she should do, rather than what she accomplished" (p. 272).

Philipp (2007) proposed that two obstacles to changing teachers' beliefs are teachers' caring and teachers' belief that teaching is telling. Cooney (1999) recommended that preservice teachers should be supported to transform their notions of overly caring for students' personal comfort levels toward caring for students' intellectual needs. Cooney implied that sometimes, pushing students to work and think outside their comfort zone can be difficult for teachers, but doing so will result in enhancing students' ability to think and analyze complex situations and concepts. Though caring for students is important, teachers do a disservice to their learning by neglecting to challenge them; this is the case of the mathematics teacher who only uses low level tasks because they believe that high level tasks are too difficult and challenging.

Teachers' beliefs have the potential to influence their instructional decisions, that is, the enacted curriculum according to the MTF. For example, Romagnano (1994) began by teaching without directly telling students what to do, allowing them to work on their own first to solve problems, whereas his collaborating teacher taught using a direct approach. Both faced difficulties as the teacher "removed much of the conceptual mathematics for the students" (Philipp, 2007, p. 280) whereas Romagnano's strategy caused students to become frustrated and disengaged from classroom activities.

According to Philipp, "[Romagnano] concluded that the differences between his and the teacher's views of mathematics, of how it is learned, and the role of the teacher in the process led to his facing the Ask Them or Tell Them dilemma" (p. 280) and the teacher did not. Here it is worth noting that teaching as telling is a common feature of low-level tasks, as such tasks require basic memorization and repetition of algorithms presented by the teacher without understanding the underlying mathematical concepts (Stein et al., 2009). Therefore, the belief that teaching as telling may lead to low quality mathematics instruction and minimizes students' opportunities to learn.

Moreover, Philipp (2007) reviewed additional research linking teachers' belief in teaching as telling and teachers' efficacy, defined as "the extent to which the teacher believes he or she has the capacity to affect student performance" (Tschannen-Moran et al., 1998, p. 202, as cited in Philipp, 2007, p. 280). Smith (1996) wrote that teachers' efficacy is challenged when reform movements advocate for teaching that contradicts their beliefs and puts their efficacy in jeopardy. According to Philipp,

> for most teachers, school mathematics is a fixed set of facts and procedures for determining answers, and the authority for school mathematics resides in the textbooks, with the teacher serving as the intermediary authority between textbooks and the students. For them, teaching requires telling, or providing clear, step-by-step demonstrations of these procedures, and students learn by listening to the teachers' demonstrations and practicing these procedures. (p. 280).

These beliefs about mathematics teaching and learning increase teachers' efficacy in the sense that it sets an attainable limit to the content knowledge required to teach and sets a low standard for teachers to attain. These beliefs assume that students' knowledge is something that must be delivered to them by their teachers, hence student learning is something that can be directly attributed to teachers' ability. This bold, yet clear picture of mathematics teaching given by Philipp and Smith challenges teachers to eradicate such beliefs and their associated practices for the betterment of student learning.

Professional Development and Teacher Change

In this section, I describe a strand of research that informs the present study by contributing two conceptual frameworks for professional development (PD) and teacher change. First, I discuss research on teacher PD that inspired the development of the frameworks and identified five features of effective PD programs. Next, I briefly describe each framework and its relevance to the present study.

Features of Effective Professional Development

The expectations and standards for student learning in school mathematics have increased since the TIMSS study in the 1990s. To guide students toward the rigorous learning objectives established in the Common Core State Standards for Mathematics (NGA & CCSSO, 2010) and by the National Council of Teachers of Mathematics (e.g., NCTM 1989, 2000, 2014, 2018), mathematics teachers must continuously enhance their knowledge and skills for teaching. Moreover, "many teachers learned to teach using a model of teaching and learning that focuses heavily on memorizing facts, without also emphasizing deeper understanding of subject knowledge" (Garet et al., 2001, p. 916). Garet and colleagues imply that there is a divide between the way that teachers have learned to teach and the expectations for student learning; this is especially true in school mathematics, where many teachers still emphasize low-cognitive processes and skills (Stein et al., 2009; Stigler & Hiebert, 1999). Therefore, PD is critical if school mathematics teachers are to enhance their knowledge and practice of teaching for higher standards (Desimone et al., 2002).

Because of the potential for PD to enhance teachers' content knowledge, beliefs, attitudes, and pedagogy, researchers have investigated both the effectiveness of PD and the components that make it effective. According to Desimone (2002), there is a consensus among researchers concerning the components of rigorous PD and that they include: (a) emphasis on content and how students learn, (b) opportunities for teachers to participate and learn actively, (c) opportunities for teachers to act and serve as leaders, (d) the length of PD programs (i.e., that they last for more than a single day or weekend), and (e) involvement from teachers who share common characteristics (i.e., that they are from the same school or teach the same grade level and content). Research investigating the Eisenhower Professional Development Program, a national program providing various forms of PD for mathematics and science teachers, contributed to and extended these ideas by investigating how they influence teachers' self-reported changes in knowledge and skills for teaching (Desimone et al., 2002; Garet et al., 2001).

The aforementioned study of the Eisenhower Professional Development Program concentrated on two components of the PD: "structural features" and "core features" (Garet et al., 2001, p. 919). Structural features included (a) the type of each PD program (e.g., study groups vs. workshops or conferences), (b) the length of each PD program, and (c) whether the PD program included teachers from the same school, department, or grade level vs. teachers from many different schools. Core features included (a) the extent to which PD activities focused on mathematics or science content; (b) the degree to which PD activities emphasized active learning for participants; and (c) the extent to which PD activities aligned with teachers' goals, state standards, and state assessments. To measure outcomes of the Eisenhower project, teachers completed Likert-type surveys to indicate self-reported increases in knowledge, skills, and changes in classroom practice (Desimone et al., 2002; Garet et al., 2001).

Garet et al. (2001) reported findings suggesting that duration of the PD and its attention to content, active learning, and connectedness to teachers' daily practice enhanced teachers' gains in both knowledge and skills. Moreover, the core features of the PD tended to make more influential differences than the type of activity (e.g., study groups vs. workshops or conferences). Finally, activities that support "coherence" (Garet et al., 2001, p. 920), such as those that are situated within teachers' daily practice, aligned with other reform efforts, and encourage teacher collaboration, appeared to influence changes in teachers' practice. Regarding the same study, Desimone et al. (2002) reported that PD focused specifically on the use of technology, higher order instruction (i.e., instruction addressing conceptual understanding), and alternative assessments (other than traditional tests and quizzes) enhanced teachers' use of each in their classrooms. Desimone and colleagues concluded that, based on their findings, "change in teaching would occur if teachers experienced consistent, high-quality professional development. But we find that most teachers do not experience such activities" (p. 102).

These findings and the need to develop, research, and refine high-quality PD supporting teacher change led to the development of five revised "core features" (Desimone, 2011, p. 69). The following features expand on the three original core features defined by Desimone et al. (2002) and Garet et al. (2001): Content focus, active learning, coherence, duration, and collective participation. Three of the (Desimone, 2011) core features stem from the three core features identified previously; the other two, coherence and duration, were formerly called structural features and retain similar meanings. The type of PD program was the only feature that did not strongly influence teachers' practice and therefore was dropped from the list. Desimone (2011) argued that the five core features are elements of effective PD that correlate with changes in teachers' knowledge and practice.

Frameworks for Professional Development and Teacher Change

In addition to establishing five core features of effective PD, Desimone (2009) acknowledged that another essential component is necessary to develop a conceptual framework for teacher PD, one that describes how PD impacts both teachers and students. In her review of empirical studies and conceptual frameworks on the subject, Desimone (2009) highlighted a model developed by Guskey (2002) that (a) informs the present study by linking teachers' attitudes and beliefs to changes in classroom practices (b) informed the development of her own framework. In the following paragraphs, I first present the Guskey (2002) model for PD and teacher change and explain its relevance to the present study. Second, I describe the Desimone (2009) model that was inspired by Guskey's and others and explain its relevance to the present study.

Guskey's (2002) model for PD and teacher change includes the following four components: (a) teacher PD, (b) change in teachers' classroom practices, (c) change in student learning, and (d) change in teachers' attitudes and beliefs. This model is linear, suggesting that teacher PD influences changes in teachers' classroom practices, changes in teachers' classroom practices influence changes in student learning, and changes in student learning influence change in teachers' attitudes and beliefs. Guskey's model differs from those that came before it because previous models supposed that changes in teachers' beliefs led to changes in classroom practices; Guskey provided his alternative model because PD programs that seek to change teachers' beliefs directly "seldom change attitudes significantly or elicit strong commitment from teachers" (p. 383). Instead, Guskey argued that teachers who experience successful implementation of new instructional practices and curricular materials are more likely to experience lasting changes in beliefs. Such teachers "believe it works because they have seen it work" (Guskey, 2002, p. 383).

The Guskey (2002) framework is relevant to the present study because it aligns with Ambrose's (2004) theory that beliefs may change based on people's experiences; moreover, the framework acknowledges Green's (2006) view that beliefs are founded on the basis of evidence (i.e., evidence of successful classroom implementation and student learning). More specifically, Guskey's model for PD and teacher change suggests that teachers' successful implementation of mathematical tasks may influence their beliefs and attitudes. According to Guskey (2002), "evidence of improvement of positive change in the learning outcomes of students generally precedes, and may be a pre-requisite to, significant changes in the attitudes and beliefs of most teachers" (p. 384). Though Guskey's framework applies to PD and not the use of tasks in general, the present study focuses on the perspectives of teachers who have been involved in task-focused PD. Therefore, there is reason to suspect that the relationship between such teachers' task use and their beliefs is bidirectional; that is, teachers' beliefs influence their use of tasks (as described previously in Chapter 2) and teachers' use of tasks may influence their beliefs as well.

Desimone's (2009) conceptual framework integrates various frameworks for PD and teacher change. For example, it builds on Guskey's (2002) framework but also includes aspects of the framework developed by Peressini et al. (2004) (reflexive interactions between teachers' practice and beliefs). Desimone's (2009) framework contains similar elements as that of Guskey, though they are presented in the following order: (a) PD, including the five core features as described in Desimone's (2011) article; (b) enhanced teacher knowledge and skills and changes in teachers' attitudes and beliefs; (c) changes in teachers' instruction, and (d) enhanced student learning. Though this model is depicted as linear, it suggests that the relationship between each successive pair of elements is reflexive; that is, (a) influences (b) and (b) influences (a), and so on.

This framework for PD and teacher change is relevant to the present study because it describes the interactive relationship between effective teacher PD and teachers' knowledge and beliefs. Each of these elements has been previously described in this review of the literature and the Desimone (2009) model interconnects them. Moreover, the model portrays the interactive relationship between teacher knowledge and beliefs and changes in teachers' instruction; it suggests that teachers who have participated in task-focused PD (i.e., the MMR and ATC) may have experienced changes in the way that they use mathematical tasks, which in turn might inspire changes in their knowledge and beliefs; it also suggests that such teachers may have experienced changes in beliefs and knowledge through their involvement in PD, leading to changes in their instruction. In summary, the Desimone (2009) framework for PD and teacher change incorporates three components that are essential to the present study: (a) teacher PD, (b) teachers' knowledge and beliefs, and (c) teachers' use of tasks and potential changes in instructional practices. Moreover, the framework illustrates how these three components of the present study are interconnected and influence one another.

The Instructional Quality Assessment Classroom Observation Tool

The purpose of this section is to discuss the instrument that was used to analyze the lessons and mathematical tasks observed in the present study. Though the primary source of data used to answer the research questions was interviews with teachers, another goal of the study was to relate teachers' responses to the use of high cognitive demand tasks. I used the IQA to analyze mathematical tasks and samples of student work.

Various researchers have pilot tested this classroom observation tool to determine its validity and reliability (Boston & Wolf, 2006; Matsumura et al., 2006) and have used it to analyze the cognitive demands of mathematical tasks (Boston, 2012; Boston & Smith, 2009); these two features support the use of the IQA and made it an appropriate tool for use in the present study. Though the IQA is not the only classroom observation tool that can be used to assess the cognitive demands of instructional mathematics tasks (e.g., Mathematical Quality of Instruction, Hill et al., 2008), the IQA is especially suitable because it tracks the progression of a mathematical task throughout a lesson, capturing the task as it appears in source materials, as it is modified in a teacher's planning, as it is set up by the teacher, and as it is implemented by the teacher and their students. The IQA rubrics tell the "story" of how a task develops throughout a lesson: "whether students engaged in thinking, reasoning, and mathematically rich discussion, and what instructional moves supported or inhibited students' engagement" (Boston & Candela, 2018, p. 431). The following section include a description the theoretical frameworks underlying the IQA and its validation and use in empirical studies. A detailed description of the instrument and its purpose for the present study is provided in Chapter 3.

Theoretical Frameworks

According to Boston and Candela (2018), the IQA is meant to measure (a) ambitious instruction, including instructional practices that support students' learning of mathematics with understanding; (b) the effectiveness of PD; (c) the implementation of new curricula, and (d) students' opportunities to learn. To measure these attributes, the IQA incorporates elements of Resnick and Hall's (1998) Principles of Learning, developed based on bodies of research in cognitive and social psychology (Boston & Wolf, 2006). The IQA is meant to address four of these principles that are "evident and observable" (Boston & Wolf, 2006, p. 5) in classrooms that support student learning: *Academic Rigor, Accountable Talk, Clear Expectations,* and *Self-Management of Learning*. Specifically, academic rigor and accountable talk are explicit in sections of the

IQA: accountable talk guided the development of rubrics for classroom discussions and academic rigor guided the development of rubrics for mathematical tasks (Boston & Candela, 2018). Each of these constructs is explained in the following paragraphs.

According to Boston and Wolf (2006), academic rigor in mathematics involves providing students with opportunities to learn mathematics for understanding; that is, students should have regular opportunities to solve problems and explain their reasoning during their engagement with rich mathematics content. Research and theory on learning mathematics for understanding and the NCTM (1989, 2000) standards influenced the focus on mathematical tasks as a means of achieving academic rigor, incorporating the work of Doyle (1983), Stein et al. (1996), and Henningsen and Stein (1997). The IQA Academic Rigor rubrics therefore address task potential and task implementation, elements of the MTF, and include criteria for analyzing tasks based on their levels of cognitive demand, similar to the TAG. Much of this research was described earlier in this chapter, so it will not be described again; however, it is important to recognize the similarities between the language used in the MTF and TAG and the language used in the IQA rubrics.

The IQA rubrics for task potential and task implementation contain examples and descriptions of task features to assist raters in scoring each item. For example, the lesson implementation section in Part 1 includes a checklist that is based on Henningsen and Stein's (1997) factors contributing to the maintenance and decline of cognitive demands (Boston & Candela, 2018). Teacher questioning (Boaler & Staples, 2008) and student opportunities to engage in mathematical discourse during whole-class discussions (Stein et al., 2009) are features of cognitively demanding tasks as implemented, and as such, these features are prevalent in the rubric for scoring the rigor of student discussions in the IQA.

The developers designed a universal scale for the rubrics (0 = low, 4 = high) so that scores of 1 and 2 align with TIMSS categories of stating concepts and using procedures and the Stein et al. (1996) categories of *memorization* and *procedures without connections*, respectively. Similarly, scores of 3 and 4 align with the TIMSS category of making connections and the Stein et al. categories of *procedures with connections* and *doing mathematics*. However, both *procedures with connections* and *doing mathematics* tasks may be rated as either a 3 or a 4 using the IQA; the difference is that a score of 4 is reserved to task for which students are required to explain their thinking. By using the same rating scale for each rubric with similar descriptors for each score level, the IQA "enables the classroom observer or interpreter of the results to develop a strong *qualitative* idea of what each score level 'looks like' in an actual classroom situation" (Boston & Wolf, 2006, p. 11).

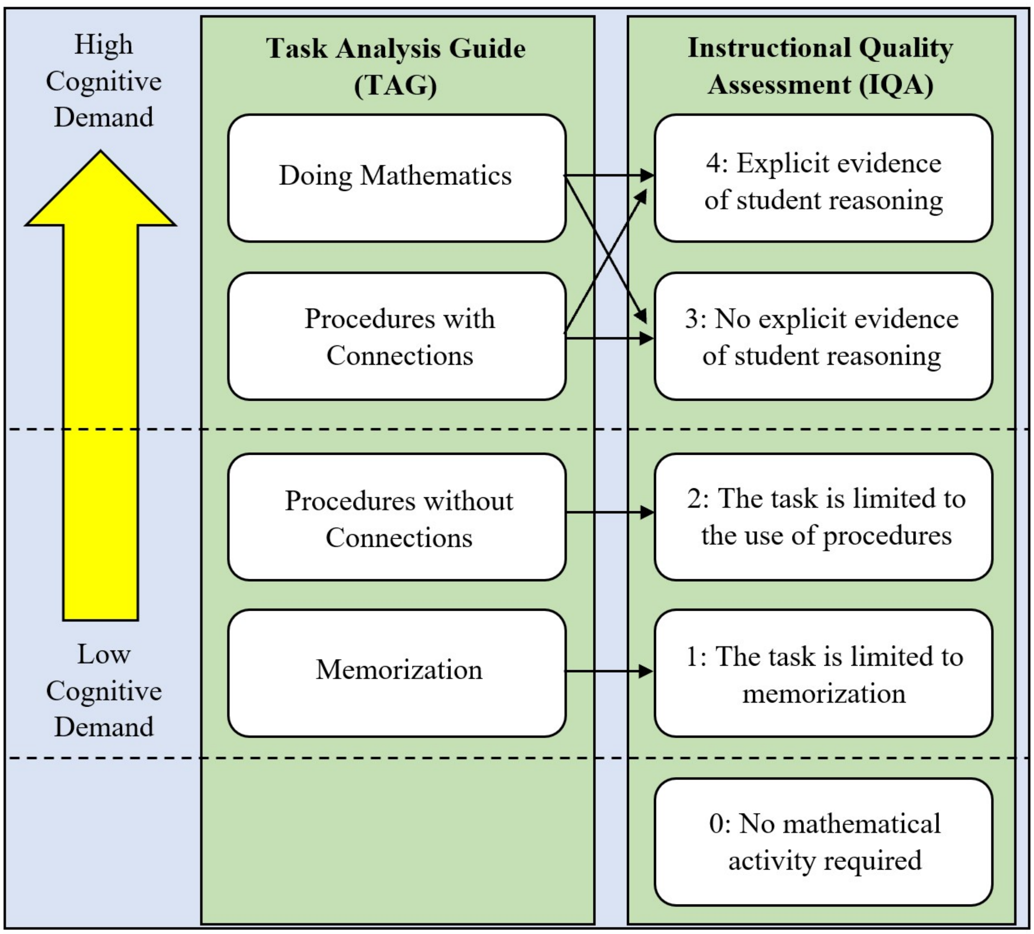
A visual comparison between the Task Analysis Guide (TAG) and the Instructional Quality Assessment (IQA) is provided in Figure 3. In this figure, the four TAG classifications and the five IQA scores (whole numbers 0–4, with brief descriptions of each as explained in the rubrics) are depicted according to the level of cognitive demand associated with each. The high cognitive demand TAG categories (*doing mathematics* and *procedures with connections*) are separated from the low cognitive demand categories (*procedures without connections, memorization, and* *nonmathematical activity*) with a dashed line segment; the same dashed segment separates the IQA scores associated with high cognitive demand tasks (scores of 3–4) from scores associated with low cognitive demand tasks (scores of 0–2). Another dashed line segment separates tasks with low cognitive demand (TAG classifications of *memorization* and *procedures without connections* and IQA scores of 1–2) from tasks that require no mathematical activity (i.e., tasks deserving IQA scores of 0).

The IQA was meant to correspond closely to the TAG, indicated with arrows pointing from each of the TAG levels to their associated IQA scores in Figure 3. Tasks within the lower three TAG categories are assigned IQA scores of 0, 1, or 2, as depicted in Figure 3. However, a *doing mathematics* task may receive an IQA score of either 3 or 4, depending on whether the task explicitly prompts for evidence of student reasoning (IQA score level 4) or not (IQA score level 3). Similarly, a *procedures with connections* task may be scored either way depending on whether explicit evidence of student reasoning is evident. This is depicted by two arrows directed from *doing mathematics* and two arrows from *procedures with connections* to the IQA scores of 3 and 4.

Accountable talk differs from traditional mathematics discussions, which follow an initiate-response-evaluate (IRE) pattern (Mehan, 1979). The IRE pattern begins with an *initiation* of a question for students to answer, usually with a specific answer in mind (e.g., what is the next step in solving this equation?). This question provokes a *response* from students, typically voiced by one student who the teacher calls on after raising their hand. Finally, the teacher *evaluates* the given student response as either “correct” or “incorrect,” after which the lesson moves on. Accountable talk substitutes *talk moves* (questions or prompts) in place of evaluation with the purpose of encouraging further student dialogue.

Figure 3

The Relationship Between the TAG and the IQA



Note. There is no formal classification for tasks requiring no mathematical activity in the TAG. Stein and colleagues (2009) discuss such tasks, but do not include them in the TAG. This portion of the TAG section in Figure 3 is left blank intentionally.

Talk moves in the IQA are categorized as either Teachers' Linking, Students' Linking, Teachers' Press, or Students' Responses; according to Boston and Candela (2018), linking includes re-voicing of others' ideas, connecting students' work and ideas, and prompting students to extend others' work and thinking. Linking moves "create opportunities for student-to-student discourse, position students as authors of ideas, and invite other students to respond, re-explain, counter, or build upon those ideas" (p. 430). Teachers' Press moves include teachers' prompting for explanations, whereas Students' Response moves are the given answers to the respective press moves. The IQA form provides a scoring rubric for each talk move with scores ranging from 0 (e.g., no class discussion or class discussion was not related to mathematics) to 4 (e.g., the teacher consistently connects speakers' contributions to each other and shows how ideas and positions shared during the discussion relate to each other).

Validity, Reliability, and Empirical Studies

The IQA was initially pilot tested in elementary schools in two similar-sized school districts in 2003 to determine its validity and reliability (Boston & Wolf, 2006). One district was involved in an ongoing reform effort, including PD focusing on the Resnick and Hall (1998) Principles of Learning whereas the other was not. The study included 6 trained graduate student raters who scored both lesson observations and student work samples. Prior research by Matsumura et al. (2002) and Clare and Aschbacher (2001) indicated that 4 work samples (2 high quality and 2 medium quality, according to each teacher) per lesson observation, rated by 2 raters each, would yield a generalizability coefficient high enough to use student work as a valid indicator of classroom practice. Matsumura and colleagues analyzed inter-rater reliability by calculating the percent of exact agreement between raters' independent scores resulting in only 50% agreement. However, 1-point agreement was excellent (95.2%). Both districts provided students with opportunities to engage in high-level tasks during mathematics instruction, however, results also indicated that the district involved in reform was significantly more effective in doing so than the other (Boston & Wolf, 2006). The study by Matsumura and colleagues replicated the findings of Henningsen and Stein (1997), among others, who have found that mathematical tasks tend to decline in cognitive demand during implementation.

A second pilot study conducted during the 2004–2005 academic year in urban middle schools aimed to (a) improve inter-rater reliability through an enhancement of the training program with more experienced raters and (b) determine the number of observations and student work samples necessary to estimate individual teachers' instructional practice (Matsumura et al., 2006). Each teacher (n = 13) participated in 2 classroom observations and submitted 4 challenging student assignments and 4 samples of student work (2 high quality and 2 medium quality, according to each teacher as in the previous pilot study). Reliability measures indicated an improvement from the previous study, with overall exact scale-point agreement from observation rubrics in mathematics at 81.8%. The exact scale-point agreement for submitted assignments was 76.3%, still considered to be moderate. Moreover, calculation of Cronbach's alpha to determine the consistency of rating scales yielded a value of α = .92 (excellent reliability), with an acceptable score being .76 or above (Matsumura et al., 2006).

Matsumura and colleagues (2006) conducted generalizability and decision studies to determine the number of observations and assignments needed to reliably measure instructional quality at the teacher level, indicating that "as few as two observations yielded a stable estimate of quality, *when teachers complied with the requirements of the data collection*" (Matsumura et al., 2006, p. 17). Prior to each observation, teachers were instructed to teach lessons including a mathematical task to work on and a whole-class discussion, essential components analyzed by the IQA rubrics. However, 2 mathematics teachers' lessons lacked these components, and their data were not included in the analyses. Matsumura and colleagues also found that collecting 4 student assignments yielded a stable estimate of instructional quality, and that reducing the number of raters from 3 to 2 did not affect the results significantly.

Though the sample size was too small to apply multi-level models, the researchers used multiple linear regression to explore the relationship between teachers' IQA ratings and student achievement, measured using the SAT-10; results indicated that, after controlling for students' background and prior achievement, IQA observation ratings significantly predicted students' scores on both the Total Math and Procedures subscales of the SAT-10 (standardized $\beta = .163$, $p < .001$ and standardized $\beta = .322$, $p < .001$, respectively). However, IQA scores for assignments only significantly predicted student achievement on the Procedures subscale (standardized $\beta = .130$, $p < .05$), echoing the connection between the tasks assigned by teachers and the student thinking that resulted (Stigler & Hiebert, 1999, 2004).

Matsumura et al.'s (2006) findings suggested that the IQA yields valid and reliable scores for measuring instructional quality and students' opportunity to learn for understanding. The reliability improved from the first pilot study after improvements to the rater-training program and using more experienced raters. However, the findings also revealed that few mathematics teachers implemented tasks that engaged students in exploring mathematical concepts and connecting them to procedures. Even when task potential was high, teachers implemented most tasks at low cognitive levels (Matsumura et al., 2006). Mean IQA scores on each rubric ranged from 1.08 to 2.65 across the 26 observations, with the majority less than 2 (the mark separating high-level and low-level cognitive demand on each of the IQA rubrics). This suggests that teachers implemented few tasks at the level of *procedures with connections* and *doing mathematics*.

Since its development and pilot testing, the IQA has been used to assess the instructional practices of mathematics teachers at the school and district levels, to identify differences in learning opportunities in large, urban school districts, to monitor teachers' instructional changes after participating in PD, and to support administrators in observing mathematics classrooms (Boston et al., 2015). Discussed previously in the review of mathematical task research, Boston and Smith (2009) used the IQA to assess the effectiveness of PD focusing on the selection and implementation of cognitively demanding tasks. The results of their study indicated that participating teachers' selection of cognitively demanding tasks increased significantly from one semester to the next and that the number of high-level tasks selected and implemented increased significantly over time. The results were not significantly different between experimental and control groups during the first semester of the study. However, the second semester yielded significant differences for task regardless of the type of curriculum that teachers used (Boston et al., 2015; Boston & Smith, 2009).

In another study to investigate the effectiveness of PD and teachers' change in instructional practices, Heyd-Metzuyanim et al. (2018) used the IQA to analyze the cognitive demand of instructional tasks and specifically focused on teachers' accountable talk during whole-class discussions. This study involved 8 middle school teachers participating in a PD program focused around the *5 Practices for Orchestrating Productive Mathematics Discussions* (Smith & Stein, 2018). Heyd-Metzuyanim and colleagues collected data in cycles, including a pre-lesson interview, a lesson recording, and a post-lesson interview, to capture potential changes in teachers' practice after engaging in the PD. The experiences of two co-teachers who succeeded in reaching the highest level of task implementation on the IQA rubrics were described in the 2018 report.

Initially, the teachers seemed to rigidly imitate what they learned during the PD without fully understanding why they were doing so and how to do so effectively; for example, both teachers made accountable talk moves during a whole-class discussion but used them to clarify students' procedures rather than to allow students to explain their mathematical thinking. However, interviews and observations 3 months after the PD revealed that both teachers were beginning to grasp ideas they had seen in the PD with greater understanding (Heyd-Metzuyanim et al., 2018). During the second pre-lesson interview, one of the teachers explained that she selected a particular task to meet students' learning needs rather than simply because the task had been used during the PD. Moreover, the teachers modified the task for their own classroom by creating their own video to introduce it while focusing on the intended mathematics content. The teachers were noticeably utilizing PD ideas in novel ways and adapting them to for use in their own classroom. During the first whole-class discussion, the co-teachers simply took a poll to see which of their students agreed or disagreed with an answer, whereas in the second lesson, they allowed one student to present his solution and facilitated a discussion to handle others' disagreements. Though it took some time, and perhaps the ability for the two teachers to co-plan and co-teach lessons following the PD, the teachers were successful in improving the cognitive demand of their tasks and engaging students in rigorous mathematical discussions (Heyd-Metzuyanim et al., 2018).

The IQA was also implemented in the Middle School Mathematics and the Institutional Setting of Teaching (MIST) project, a 4-year investigation of mathematics teaching and student achievement in 114 middle school classrooms within 4 urban districts (Boston & Wilhelm, 2017; Jackson et al., 2013). Boston and Wilhelm reported on the data from the first year of the project, whereas Jackson and colleagues' article contains results from years 3–4. During the first year of the MIST project, the research goals were to investigate the rigor of instructional tasks, opportunities for students to engage in mathematical discussions, differences between the four districts involved in the study, and how the data compared to previous studies using classroom observations (Boston & Wilhelm, 2017).

Consistent with previous studies, the MIST team observed and coded 2 (consecutive, when possible) mathematics lessons using the IQA rubrics for academic rigor and accountable talk. One of the key findings during the first year of the MIST project was that task potential “sets the ceiling” (Boston & Wilhelm, 2017, p. 852) for both task implementation and the student discussion that may follow. That is, IQA scores for task implementation and student discussions rarely exceeded scores for task potential. This result echoes the findings from both the TIMSS (Stigler & Hiebert, 1999) and QUASAR (Henningsen & Stein, 1997; Stein et al., 1996, 2009) studies, indicating that the cognitive demand of a task tends to decline from setup to implementation. Moreover, IQA task implementation scores were significantly lower than task potential scores in Year 1 of the MIST study (Boston & Wilhelm, 2017). Another interesting finding was that task implementation scores differed significantly across districts. The researchers suggested that the teachers’ experience, long-term use of *Standards*-based curricula, and involvement in PD may have contributed to such differences (Boston & Wilhelm, 2017).

Data from Years 3 and 4 of the MIST study were analyzed to explore teachers’ task setup and the relationship between task setup and students’ opportunity to learn during whole-class discussions, including observations of 165 teachers’ instruction over 2-day periods (Jackson et al., 2013). In addition to the original IQA rubrics, the researchers developed a set of accompanying rubrics to assess task setup, including rubrics for (a) contextual (real-world) features of task scenarios, (b) mathematical relationships evident in the task statement, (c) the maintenance of the cognitive demand of the task, (d) the cognitive demand of the task at the end of the setup as students begin to work on the task, and (e) the percentage of students who participated in the setup discussion. Using both sets of rubrics to analyze classroom observations, Jackson and colleagues found that the level of attention to mathematical relationships in task setup was positively related to the quality of the whole-class discussions that followed, regardless of whether there was a contextual or real-world element to the task. The results also showed that discussions were of higher quality when students and teachers attended to contextual features and mathematical relationships during task setup and the cognitive demand of the task was maintained, though such conditions were rarely met. Similar to previous studies, the cognitive demand of tasks tended to decrease from setup to implementation, occurring in more than 60% of observed lessons (Jackson et al., 2013).

Boston and Candela (2018) examined mathematics teachers' instructional practices and highlighted ways in which the IQA can serve as a tool to enhance them. The authors observed and rated 3 lessons, 2 of which received low scores on the IQA (each item scored as either 1 or 2) and the third more closely resembling ambitious instruction (items scored between 2 and 4 with a mode of 3). A low-scoring geometry lesson was primarily teacher-directed, with few students providing short answers to questions and completing small tasks. The second low-scoring lesson involved the procedure for multiplying whole numbers by fractions; during this lesson, the teacher gave detailed instructions to show students how to complete each task, with little time spent on student discussions.

The third, more ambitious lesson, engaged students in investigating multiplication problems where a factor was either doubled or halved. This lesson began with a class discussion to activate students' prior knowledge, followed by a task in which the teacher played the role of a facilitator by providing feedback and scaffolding to individuals and groups. The teacher of this lesson provided students with opportunities to make mathematical connections and generalizations, allowing them to engage in higher-level thinking about products and factors (Boston & Candela, 2018). This study highlighted the ways in which mathematics teachers can use the IQA as a tool to identify and reflect on aspects of their own practice.

The Expanded Instructional Quality Assessment and the Task Setup Rubrics

One outcome of the MIST project was the development of task setup rubrics to accompany the standard IQA rubrics; the additional rubrics are known as the Expanded Instructional Quality Assessment (EIQA) Task Setup Rubrics. Interest in task setup arose during the MIST study as researchers sought to identify ways that teachers supported *all* students to engage with mathematical tasks meaningfully. Task setup was of particular importance because it occurs just prior to task implementation in the MTF and influences which students are able to engage in solving a task and how they might go about solving it (Jackson et al., 2012, 2013). By watching and analyzing video recordings of middle-grades mathematics instruction, Jackson and colleagues (2013) noticed that the work of both teachers and students during task implementation was determined by the way in which the task was set up. For example, "when students are not supported to understand key aspects of the task statement, teachers often spend the next phase of instruction reintroducing the task to individuals or groups of students while others begin to solve the task" (p. 649).

Through a qualitative analysis of 40 video-recorded episodes of middle-grades mathematics teachers' instruction collected during Year 1 of the MIST project, Jackson and colleagues (2013) identified four aspects of high cognitive demand task setup that tended to support students' engagement: (a) important contextual features (i.e., real-world aspects) of the task scenario were discussed, (b) important mathematical relationships represented in the task were discussed, (c) a common language was developed to discuss both contextual features and mathematical relationships, and (d) the cognitive demand of the task was maintained throughout task setup. The contextual features of a task are critical because some students may be unfamiliar with the real-world context of a task if the context is outside their knowledge or experience. For example, consider a task that has students maximize the volume of a water trough for animals in a barn using a fixed amount of materials. Students who have never experienced animals on a farm may not have a mental picture of what a water trough looks like and may lack the prior knowledge to get started on the task without a teacher's support (or the support of other students). To enhance students' engagement with the task and maintain the cognitive demand, a teacher implementing this task may begin with a whole-class discussion or provide pictorial examples to access students' background knowledge and experiences (Jackson et al., 2012, 2013).

Similarly, discussing key mathematical relationships during task setup can support high-level task implementation (Jackson et al., 2012, 2013). Consider a task that involves comparing various cellphone data plans where some plans charge a fixed cost each month and others cost various rates depending on the data usage each month. To determine which plan is the most cost-efficient, it is crucial for students to understand the differences between the payment options for each plan and how the monthly cost may increase with the data usage. This does not mean that the teacher simplifies the work of students or does the mathematical thinking and work for them; rather, it means that the teacher supports students in understanding the mathematical quantities and relationships in the task statement and poses questions to make sure that this understanding is shared by all students.

The third aspect of high-level task setups identified by Jackson et al. (2013) was the development of a common language to describe key contextual features and mathematical relationships. This requires the teacher to illicit responses from multiple students and ask questions that prompt students to use relevant real-world and mathematical language (Jackson et al., 2013). Teachers must "build on student contributions and both support and press students to develop common language to describe key features of the task" (Jackson et al., 2012, p. 28). Developing a common language allows students to be able to communicate more effectively with each other and with the teacher.

The fourth aspect of high-level task setups, maintaining the cognitive demand of the task, has been described in previous sections. However, it is worth reiterating that a high-level task setup is one in which the teacher does not do the mathematical work and thinking for the students and maintains the cognitive demand of the task (Stein et al., 2009). Maintaining the cognitive demand of the task means that the setup discussion of contextual features and mathematical relationships is student- rather than teacher-dominated and the teacher allows students to explain how these aspects are evident in the task statement.

The four aspects of high-level task setups described in the previous paragraphs provide the theoretical and empirical foundation for the EIQA Task Setup rubrics. The rubrics were meant to be compatible with the standard IQA and hence are measured on a similar 0–4 scale, closely aligning with the Stein et al. (2009) TAG. Rubric 1: Contextual Features (CF) measures the level at which students are supported in developing a shared understanding of the contextual, or real-world, features of a task (if a task has a real-world, problem-solving scenario). Similarly, Rubric 2: Mathematical Relationships (MR) measures the level at which students are supported in developing a shared understanding of the mathematical relationships and ideas represented in a task.

For either rubric, a score of NS indicates that there was no whole class discussion of the task prior to students' engagement, whereas a score of N/A on each rubric indicates that the task does not have a problem-solving (real-world) scenario or is not mathematical in nature, respectively. Scores of 0 suggest that there was no attempt to discuss contextual features or mathematical relationships, and scores of 1, 2, 3, and 4 consistently increase in rigor from superficial student engagement to students consistently making connections between ideas. Both rubrics emphasize the teacher's use of accountable talk moves, however, a subtle difference is that the MR rubric requires the presence of a *strong* accountable talk move on the part of the teacher or the students (Jackson et al., 2013).

Jackson et al. (2013) used the EIQA rubrics to code 460 video-recorded teaching episodes from Years 3 and 4 of the MIST study. Of the 460 lessons, 58% (*n* = 267) involved a problem-solving scenario and warranted the use of both rubrics whereas the other 42% (*n* = 193) lessons did not, and therefore were not coded using the CF rubric. Using regression models, the researchers found a positive relationship between the quality of the attention to mathematical relationships and the quality of the concluding whole-class discussions at the end of a lesson. Even in lessons without a problem-solving scenario, a positive relationship existed between the quality of the task setup and the quality of the concluding whole-class discussion. However, in lessons with problem-solving scenarios, teachers more frequently addressed mathematical relationships than contextual features and did so at higher levels. These results suggest that when teachers and students discuss the contextual features and mathematical relationships and maintain the cognitive demand during the setup of a task, the concluding whole-class discussions may be of higher quality (Jackson et al., 2013). Hence, the setup of a task is crucial in how the task is implemented and whether the cognitive demand is maintained.

I chose to use the two EIQA rubrics in addition to the standard IQA because they provided me with two aspects of task setup to analyze in further detail: teachers' attention to contextual features and their attention to mathematical relationships. The addition of the two EIQA rubrics strengthened my analysis of tasks at the set-up phase and added depth to the study. An additional factor that led to the inclusion of the two rubrics was that the Ohio MMR course was designed to include tasks with real-world connections and applications; the course is meant to emphasize the use of mathematical models to simulate real-world problem-solving scenarios. As such, many of the tasks that teachers used in this study could be analyzed using both rubrics. Analyzing the rigor of MMR teachers' task setup in greater depth has implications for the MMR pilot course and similar PD incentives focusing on high-level task use.

Chapter Summary

Mathematical task research expanded from the research of Doyle (1983, 1988), who developed academic tasks as a construct and categorized them according to the type and amount of work required by students to complete them. Stein and colleagues' (1996) TAG expanded on Doyle's research in mathematics education specifically by classifying tasks into four types based on their level of cognitive demand: *memorization* tasks, *procedures without connections* tasks, *procedures with connections* tasks, and *doing mathematics* tasks. Stein et al. (1996) additionally developed the MTF to follow the progression of a task from written source materials to implementation in the classroom.

Research in mathematics education has provided evidence suggesting that cognitive and affective factors, such as mathematics teacher beliefs and knowledge, influence the selection, planning, and implementation of instructional tasks (e.g., Charalambous, 2010; Collopy, 2003; Hill, Blunk, et al., 2008; Philipp, 2007), providing motivation to explore such factors in further detail. Various components of MKT (Hill, Ball, & Schillings, 2008) and beliefs about tasks (Collopy, 2003; Philipp, 2007), beliefs about students (Philipp, 2007; Romagnano, 1994), and beliefs about mathematics teaching and learning (Beswick, 2011) have each manifested in teachers' use of tasks at each phase. Finally, this chapter concluded with a description of the IQA, an instrument that yields valid and reliable scores to measure the quality of instruction (Boston & Wolf, 2006; Matsumura et al., 2006) and the cognitive demands of instructional tasks (Boston, 2012; Boston & Smith, 2009).

Chapter 3: Methodology

In this chapter, I present the research methods for the present study. I begin by describing and justifying the research design. Next, I provide background details by explaining the research context and participants from which I collected data. To set the stage for data sources and collection procedures, I include a brief summary of my IRB protocol and the IQA research instrument that I used to analyze mathematical tasks. After discussing data sources and the data collection procedures that I used, I describe the data coding and analysis processes employed. To conclude the chapter, I address issues that pertain to the study as a whole, including my positionality and credibility, transferability, and trustworthiness—constructs in qualitative research that are analogous to validity and reliability in quantitative research.

Research Design

Because the goal of the present study was to explore how high school mathematics teachers perceive and reflect on the use of mathematical tasks, I used primarily qualitative research methods to collect and analyze data (Rubin & Rubin, 2012). The research questions address the perspectives of high school mathematics teachers, which cannot be deeply understood through quantitative measures. Qualitative methods captured depth of understanding rather than breadth (Rubin & Rubin, 2012), allowing for research participants to explain their thinking and rationales for selecting, planning, setting up, and implementing mathematical tasks. Qualitative methods are inductive and naturalistic in the sense that the researcher might not influence the phenomena that are studied (Patton, 2015); in the present study, high school mathematics teachers' use of mathematical tasks unfolded naturally without my interference, whereas my goal was to learn from what teachers did and thought throughout the process. Within qualitative methodology one of the primary methods is interviews; this method was used to collect data to address the research questions, supported by the collection and analysis of documents (mathematical tasks and samples of student work) and remote observations.

Interviews

According to Brinkman and Kvale (2015), an interview is a conversation with a purpose, a discussion between the researcher and participant(s) with the purpose of constructing and uncovering knowledge. Interviews are especially useful to study processes that are not externally visible (Rubin & Rubin, 2012), in this case, to study the motivating forces that drive high school mathematics teachers' use of tasks at each of the four task phases. Through interviews, teachers were provided opportunities to explain why they selected particular tasks, why they modified them during their planning, how they set the tasks up for their students, and why they made particular instructional moves during task implementation. Interviews were necessary because the reasons for such actions could not be discerned through observations or documents alone. Additionally, qualitative interviews allowed for participants to express multiple views, including those that were not necessarily expressed in the literature (Brinkman & Kvale, 2015).

Qualitative interviews consist of main questions, probes, and follow-up questions (Rubin & Rubin, 2012), each of which I used for the present study. Main questions were designed to answer the research questions (e.g., why did you choose to use this task with your students?), whereas probes pressed interviewees to provide further examples and details (e.g., how does this task address math concepts?) and follow-up questions asked participants to elaborate on ideas in more depth (e.g., could you explain what you mean by…?). The main questions that I constructed for interviews were based on my knowledge and experience with mathematical tasks and based on the literature (Rubin & Rubin, 2012); for example, I categorized questions into four groups, each pertaining to a single task phase. Of the three types of interviews, I chose to use semistructured interviews (Glesne, 2016; Merriam, 2009) in the present study because semistructured interviews provided structure through a predetermined set of questions but also allowed me to ask follow-up questions as needed. Neither structured nor unstructured interviews would have offered the same balance of support and flexibility. Tentative pre- and post-observation interview questions are presented in Appendices C and D, respectively.

Documents

Interviews with high school mathematics teachers were situated within the context of the mathematical tasks they selected, planned, set up, and implemented. Moreover, one of the goals of the present study was to explore such teachers' abilities to analyze tasks and how they reflected on the potential changes between the four task phases. Therefore, I collected and analyzed mathematical tasks at each of the four task phases, documents that enhanced interviews by providing additional insights and potential interview questions (Patton, 2015). For example, access to the mathematical task used by each teacher allowed me to develop additional interview questions pertaining to the mathematics content and the cognitive demand of each task. For example, I was able to ask questions such as "What might you do if students struggle to identify a reasonable pair of independent and dependent variables?" By collecting the same task at each of the four task phases, I was able to ask questions to investigate why teachers made particular modifications to the task from one phase to the next and how the changes influenced the cognitive demand of the task, such as, "Can you tell me about the changes you made to number 2 [on the handout]?" and "What led you to make this change?" Mathematical task documents provided a window into teachers' perspectives and experiences (Patton, 2015) as they engaged their students and reflected on their instruction.

Observations

Mathematical tasks as they appear in source materials and in teachers' lesson plans served as documents that could be collected and analyzed in digital form. During observations of teachers' mathematics instruction, I wrote detailed field notes to describe what I observed, what I wondered about it, and questions that I planned to ask later during interviews (Glesne, 2016; Patton, 2015). For example, I made written notes indicating when task setup ended and when task implementation began to guide my subsequent IQA analyses. Both observations and field notes focused on the mathematical tasks that were set up and implemented and on the evidence needed to analyze their cognitive demands using the IQA rubrics. I documented evidence of students' thinking and reasoning by recording instances when they explained their mathematical thinking verbally. I specifically documented talk-moves made between teachers and students because these statements are emphasized in the rubrics.

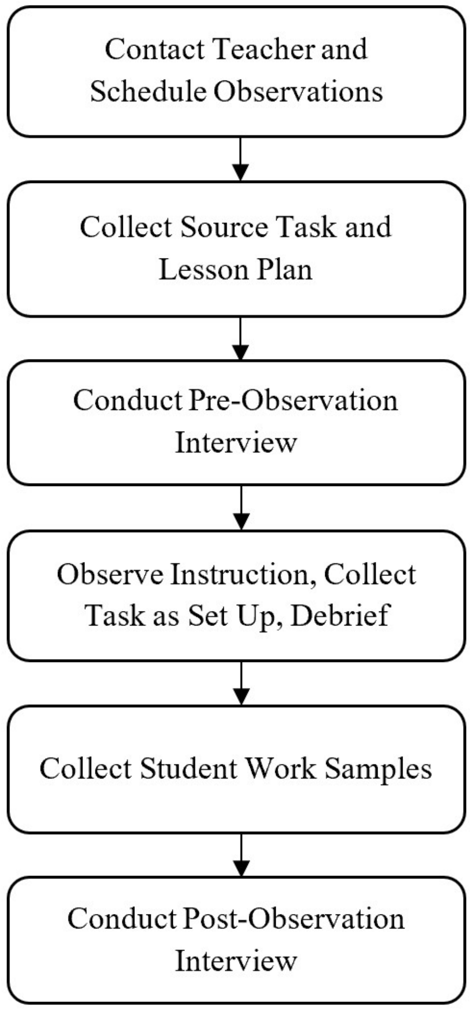
Data Sources and Collection Procedures

To answer the research questions, I collected 3 types of qualitative data: interviews, observations, and documents (Glesne, 2016; Patton, 2015; Rubin & Rubin, 2012). Interviews served as the main source of data to answer the research questions, focusing on high school mathematics teachers’ perspectives to determine (a) reasons that they attributed to the selection, planning, set up, and implementation of instructional tasks, (b) reasons that they attributed to changes in task cognitive demands between phases, and (c) reasons that they attributed to the high cognitive demand of tasks that were identified as low cognitive demand according to the IQA. I observed and analyzed instruction to determine the cognitive demand of implemented tasks and used observational field notes to develop follow-up interview questions. I collected and analyzed mathematical tasks to determine their cognitive demand at each of the four task phases (selected, planned, set up, and implemented) and to provide context for interviews with teachers. I also collected student work samples to augment the analysis of observations and to determine the cognitive demand of implemented tasks.

Figure 4 provides a pictorial outline of the data collection procedures described in the previous paragraphs. This progression of data collection procedures occurred once for each teacher. The following paragraphs include a description of the data collection process that took place with each participant in chronological order.

Figure 4

Stages in the Data Collection Process



Note. Data collection follows the progression of a single mathematical task through the four task phases: as selected, as planned, as set up, and as implemented. A single task was collected from each participating teacher, serving as the focus of observation and both interviews. Two observations were conducted to assess the full implementation of a task across two class meetings. Each process depicted in Figure 4 occurred once for each teacher, except that a debrief was not possible in every instance.

Mathematical Tasks Selected from Source Materials and as Planned

The purpose of the present study was to follow the progression of mathematical tasks from source materials to implementation and learn about teachers' perspectives as they go through and reflect on the process. After sending emails to potential research participants and obtaining their consent to conduct research, I scheduled 2 consecutive observations with each. Consecutive, in this context, refers to consecutive days of instruction with the same group of students. Instances where teachers met with their students every other day (e.g., Monday and Wednesday block scheduling) were considered as consecutive observation days even though instruction did not occur on consecutive days of the week.

I required teachers to implement what they considered to be a high cognitive demand task during the observations. Prior to each pair of observations, I asked teachers to submit a mathematical task from the lessons that they believed was a high cognitive demand task, yielding a total of 9 tasks to be analyzed. Such tasks were considered by teachers to be the *main task* for each observed lesson, either the task consuming the largest amount of class time or the task designated by teachers as the main instructional task for the lesson (Boston & Smith, 2009). I limited data collection and analysis to 1 task per teacher, even in instances where a second task was implemented over the 2 observations, so that I could concentrate on the progression of the single task during interviews. I anticipated, and later confirmed, that a 1-hr interview was not enough time to discuss the selection and planning of multiple tasks in depth; similarly, 1 hour was not enough time to discuss the setup and implementation of more than a single task.

I collected each main task as it appeared in source materials and as it appeared in teachers' lesson plans so that I could analyze tasks at both the selected and planned phases. Tasks from source materials were located in curricular materials, such as textbooks, found using the internet, or created by the teachers; I accepted main tasks in any of these forms. I also requested the full lesson plans for each observed lesson to support my analysis of tasks at the planned phase. The task that each teacher selected and the adjustments they made when planning are described in Chapter 4.

Pre-Observation Interviews

After collecting evidence of selected and planned tasks, I scheduled a *pre-observation* interview which took place prior to the observation days (for a total of 9 pre-observation interviews, 1 for each teacher) focusing on the tasks selected from source materials and included in teachers' lesson plans. All interviews were conducted remotely, using Zoom software for digital communication to avoid the health risks of meeting face-to-face due to the COVID-19 pandemic. Remote interviews were each scheduled for 1 hr at times convenient with research participants.

Prior to conducting the interviews, I developed an interview protocol (Glesne, 2016; Patton, 2015) for pre- and post-observation interviews consisting of possible questions that I planned to ask each participant, shown in Appendices C and D, respectively. Introductory questions asked participants to explain why they selected a task, why they made changes to a task when planning, and why they made instructional decisions when setting up and implementing the task with students. Follow-up questions delved deeper and probed to identify the underlying reasons why such decisions were made.

During these one-on-one interviews, I asked teachers to explain the thought processes and rationales for their decision-making when they selected and planned each main task. For example, I asked questions such as, "Why did you choose to use this task with your students," with the follow-up question, "How does this task address student engagement?" to determine how tasks met teachers' goals for student engagement. I also had teachers categorize each main task using the TAG (e.g., as *memorization*, *procedures without connections*, *procedures with connections*, or *doing mathematics*); after teachers rated the tasks that they selected from source materials and the tasks as planned, I asked interview questions to illicit justifications for each rating and explanations for changes (or lack thereof) in the cognitive demand between them.

After interviewing each participant and transcribing the interview into text, I sent the interview transcript to the participant to verify its accuracy. By conducting such a member check, I ensured that the interview data truly captured my participants' perspectives (Glesne, 2016; Merriam, 2009). This process was done for both pre- and post-observation interviews, though it was done separately because each type of interview typically occurred on a different day.

Mathematical Tasks as Set Up and as Implemented

I then observed the setup and implementation of each task with another trained IQA rater, another graduate student in mathematics education at Ohio University. All observations were done remotely, using software such as Zoom or Google Classroom, to avoid face-to-face contact with teachers and students during the COVID-19 pandemic. Rather than video- or audio-recording each lesson, both observers took written field notes as they observed instruction remotely to emulate the methods of previous studies using the IQA (e.g., Boston & Wilhelm, 2017; Jackson et al., 2013). Observations served to capture evidence of the setup and implementation of each main task. Recall that task setup was defined as the task that was initially presented to students in class, including directions given orally and in writing (Stein et al., 1996; 2009). Therefore, both IQA raters included all directions given to students in their observational field notes and collected digital copies of each main task as presented to students (for a total of 9 set up tasks, 1 for each teacher).

Both IQA raters transcribed lessons in detail in their observational field notes (Glesne, 2016; Patton, 2015) with specific attention to the nature of mathematical tasks as they unfolded and the interactions between students and teachers; as the observations were done remotely, this included documenting what could be seen and heard via webcams. Throughout the data collection process, both observers did their best to accommodate for teachers' and students' needs and to limit interference with their regular classroom procedures. Participating teachers taught in a variety of face-to-face, remote, and hybrid environments, each following the guidelines and policies designated by their respective school district. Some face-to-face classroom setups resembled life before COVID-19, with the exception of masks, social distancing, and additional sanitation materials. Some classes met remotely via Google Classroom or Zoom, where teachers used breakout rooms to put students into groups and facilitate online student collaboration.

Regardless of the teaching and learning situation, collecting data during the COVID-19 pandemic was not ideal, though it did provide additional data that might not have been obtained otherwise. In some instances, student-to-student communication was difficult, if not impossible, to hear and transcribe based on where students were seated and where the webcam was placed in the room. Fortunately, many teachers revoiced their students' contributions during whole-class discussions and some teachers used multiple devices to help capture audio and video data (e.g., using cellphones to capture student talk or the use of two webcams to view the classroom from various angles). Under normal circumstances, both researchers would have been able to walk around each classroom and observe students as they worked and communicated among themselves. This is a limitation to the present study, however, analyzing student work samples using the IQA provided data to determine the level of students' thinking and reasoning in addition to what was obtainable through remote observations.

Both raters used the IQA to assist in recording lesson details and as a guideline for written notes, including accountable talk moves (Boston, 2012) made by the teacher and students. In accordance with the IQA, both raters also recorded questions asked by the teacher and students, including question types such as probing and exploring mathematical meanings and relationships. Linking moves, teachers' asking moves, and students' providing moves (Boston, 2012) served as an additional focus for written descriptions during lesson observations. Immediately following each observation, I debriefed with each participating teacher one-on-one if possible and asked them to explain how they felt that their tasks unfolded and why. Most teachers were able to debrief briefly after each lesson, but some were unable to because they needed time to prepare for the following class. Debriefs served as short post-observation interviews with teachers immediately following task implementation so that I could ask questions and take notes of teachers' responses with the lessons still fresh in their minds. For each observed lesson, I created a timeline of lesson activities (Boston, 2006) and scored task setup and implementation with the help of my co-investigator. The data analysis procedures are described in a later part of this chapter.

After observing the implementation of each main task, I requested that teachers submit student work samples from each task to support my analysis. Research participants submitted at least 6 samples of student work, if possible (Boston, 2012; Matsumura et al., 2006). Each submitted digital copies of students' work meeting the following criteria, whenever possible: 2 samples displaying strong student work, 2 samples displaying average student work, and 2 samples of the teacher's choice. The decision to collect samples this way was based on previous research suggesting that at least 4 samples (2 high quality, 2 medium quality) are required to represent valid indicators of classroom practice using the IQA (Boston & Wolf, 2006; Matsumura et al., 2006).

Two teachers chose to submit all copies of student work (more than 6 samples each); in these instances, I determined the 6 student work samples to be analyzed according to the criteria discussed previously. As I explain in the Data Coding and

Analysis section of this chapter, I then used IQA Rubric AR2: Implementation of the Task to rate observations and student work samples with my co-investigator, yielding IQA scores for (a) task implementation and (b) students' work. To keep students' names confidential, participating teachers concealed or removed their names before submitting copies of their work. A full synopsis of the collected student work is reported in Chapter 4.

Post-Observation Interviews

Finally, I conducted a remote, one-on-one post-observation interview with each teacher following the observations to discuss the setup and implementation of each main task (for a total of 9 post-observation interviews). Due to teachers' busy schedules, I was unable to conduct these interviews immediately following task implementation, hence the need for short debriefs following each lesson. However, I attempted to schedule post-observation interviews as promptly as possible so that research participants could recall each lesson more clearly. During post-observation interviews, I invited participants to explain how their tasks progressed and to explain their instructional decisions based on the field notes I took while observing. For example, I asked questions such as, "Can you recall an instance where student(s) struggled with this task?" and followed up with, "How did you respond in this instance?" to explore what motivated teachers' actions when responding to student difficulties. Additionally, teachers rated task setup and implementation using the TAG and provided reasoning for changes (or lack thereof) in the cognitive demands of tasks from planning to setup and from setup to implementation.

Though I used the IQA to analyze instructional tasks, I had teachers use the TAG because of their prior experience with it in PD and because they were more familiar with its use and purpose than the IQA. It is unlikely that any of the research participants were familiar with the IQA and would have required additional training to implement the IQA rubrics effectively. Moreover, the IQA instrument was also meant to be used by observers rather than the teacher who taught the lesson (Boston, 2012; Jackson et al., 2013). However, the TAG adequately served as a tool that research participants could use to measure the cognitive demands of tasks at each phase.

During both pre- and post-observation interviews, participating teachers had the potential to identify a task as high in cognitive demand (i.e., *procedures with connections* and *doing mathematics*) using the TAG whereas I might have identified it as low cognitive demand using the IQA (i.e., scores of 0–2). However, I chose not to reveal IQA scores nor my analyses of instructional tasks so that I did not influence teachers' task analysis and their own explanations. I wanted the interview data to be authentic and uninfluenced by my own perspectives. A second type of mismatch could have occurred, where a teacher identifies a task as low in cognitive demand whereas I identify it as high cognitive demand; however, no such cases occurred in the present study. Because I did not reveal the IQA scores for teachers' tasks, I instead compared teachers' analysis of their tasks to my own to determine potential reasons for each mismatch that happened. A description of the three mismatches in task analysis is presented in Chapter 4 and discussed further in Chapter 5.

The Instructional Quality Assessment Research Instrument

The IQA is a classroom observation form that is filled out by classroom observers during and immediately following a lesson. Detailed field notes are to be taken during the lesson itself and attached to the form, serving as evidence to support the ratings (scores) given using the IQA rubrics. The instrument is divided into three parts:

1. Documents needed during the observation,
2. The IQA mathematics rubrics, and
3. The scoring sheet.

Part 1 provides guidelines for what to look for when observing the enacted lesson, including *accountable talk* and *academic rigor,* two theoretical constructs that were discussed in Chapter 2. The Accountable Talk Function Reference List includes "talk moves" (Boston & Candela, 2018, p. 430), verbal actions which either provide accountability to the learning community (e.g., keeping students together to follow complex thinking) or provide accountability to knowledge and rigorous thinking (e.g., asking students to explain their reasoning). Part 1 also includes a lesson implementation checklist that helps the observer to determine whether the lesson provided students with opportunities to engage in high-level thinking and reasoning, specifically addressing the quality of the class discussion following implementation of a mathematical task and the types of questions asked throughout (e.g., probing questions, questions exploring mathematical relationships, and so on).

Part 2 of the IQA contains five rubrics for academic rigor and five rubrics for accountable talk that were described in Chapter 2. Each IQA rater observes instruction, takes detailed field notes, and scores each item on the rubric based on what was observed. Discussed in more detail in the Data Coding and Analysis section, both IQA raters scored tasks from source materials, task planning, and task setup using Instructional Quality Assessment Academic Rigor Rubric 1: Potential of the Task (IQA Rubric AR1) and scored task implementation using Instructional Quality Assessment Academic Rigor Rubric 2: Implementation of the Task (IQA Rubric AR2).

Part 3 is the scoring sheet for which 2 raters provide their scores to each rubric. Scores are given as single whole numbers between 0 and 4, inclusive, with higher scores indicating stronger evidence of ambitions mathematics instruction (Boston, 2012). The IQA requires completion by 2 trained raters for each lesson to maintain validity and reliability of the rubrics (Boston & Wolf, 2006; Matsumura et al., 2006). Researchers utilizing the IQA in previous studies have required raters to reach 80% exact agreement prior to coding data as an additional reliability measure (Boston & Wilhelm, 2017; Wilhelm & Kim, 2015), though I implemented the rubrics with a co-observer by reaching a consensus through discussion. The full IQA rater form for lesson observations is included in Appendix A. I also used two EIQA rubrics, Contextual Features (Rubric 1), referred to as CF, and Mathematical Relationships (Rubric 2), referred to as MR; these rubrics were meant to analyze elements of whole-class discussions that occur during task setup. Both setup rubrics are presented in Appendix B.

Context and Research Participants

I purposefully selected research participants meeting the following criteria: first, they were high school mathematics teachers in the state of Ohio. As discussed in Chapter

1, I planned to select teachers from Ohio to make data collection manageable, though this was no longer necessary due to the use of only remote data collection procedures. I decided to keep this criterion for selection because I wanted to explore the perspectives of high school mathematics teachers involved in ATC programs and the Ohio MMR pilot course. Second, research participants must have engaged in *task-focused* PD including the Stein et al. (2009) MTF and TAG; this was so that I could recruit teachers who were more familiar with mathematical task terminology (e.g., words such as tasks and cognitive demand), providing a shared understanding of vocabulary for interviews.

Third, I required that each teacher planned to use what they considered to be a high cognitive demand task that I could observe and collect student work samples from. The focus of the present study was to investigate high school mathematics teachers' perceptions regarding *high cognitive demand* tasks; however, the tasks used by teachers did not always align with high cognitive demand tasks according to the IQA rubrics (i.e., scores of 3–4). In such instances, I asked interview questions to illicit teachers' justifications for why they believed that the cognitive demand of the tasks was high to delve deeper into their perspectives.

Purposeful selection of research participants meeting the three criteria served to provide *information rich* cases to study in depth (Patton, 2015). Qualitative research tends to consider a small number of such cases but determining sample size can be difficult. Ideally, qualitative researchers continue to recruit new research participants until data saturation is met. Guest et al. (2006) define data saturation as "the point at which no new information or themes are observed in the data" (p. 59). Through their coding and analysis of interview data, Guest et al. found that data saturation was reached after 12 interviews and that additional interviews rarely produced new codes. Therefore, I aimed to recruit 12 high school mathematics teachers to each participate in a pre-observation and a post-observation interview, using Guest et al.'s number as a guideline.

Though I sent an initial email message to more than 50 teachers and received letters of support from 12 principals, my sample includes only 9 teachers. One teacher withdrew prior to data collection for health reasons and one teacher's district abruptly changed to asynchronous, remote instruction, making classroom observations impossible. With a letter of support from this teacher's principal, I sent an email to the teacher asking if he would still like to participate in the interviews but received no response after sending several follow-up messages. The third teacher for whom I received a principal's letter of support agreed to participate in the study via email, but after many weeks of follow-up emails, failed to respond and schedule observation dates. I chose to continue with data collection and analysis with a sample of 9 teachers because the data provided enough information to sufficiently answer the research questions. However, I remained flexible and was willing to recruit additional teachers if I had found it necessary.

To recruit research participants who had engaged in PD including the Stein et al. (2009) MTF and TAG, I contacted (a) pilot teachers of the Ohio Mathematical Modeling and Reasoning (MMR) course during the 2010–2021 academic year and (b) past participants of Advanced Teacher Capacity (ATC) PD programs hosted at Ohio University. Among the teachers I contacted, I prioritized those whose principals or superintendents gave their consent to conduct research and those who responded to my invitation promptly. This convenience sample may have introduced potential bias to the present study (Patton, 2015; Glesne, 2016), as high school mathematics teachers in other locations and those who did not participate, but were eligible to do so, may hold different perspectives concerning mathematical tasks. The following subsections include brief descriptions of the Ohio MMR pilot course and the ATC PD programs and justification for the inclusion of each in the present study.

The Ohio Mathematical Modeling and Reasoning Pilot Course

The Ohio Mathematical Modeling and Reasoning (MMR) course is a student-centered high school mathematics course designed to serve as an alternative to Precalculus for seniors who do not intend to pursue STEM pathways. The course is organized around nine themes: Problem Solving, Number and Quantity, Functions–Part 1 (linear functions), Functions–Part 2 (quadratic, exponential, and power functions), Geometry, Statistics, Probability, Applications of Number and Quantity and Statistics, and a Wrap Up theme at the end of the course. One purpose of the Ohio MMR course is to serve as a transition course that prepares high school seniors for remediation-free college mathematics. The course is intended to be taught rigorously using effective teaching practices identified by mathematics education research and professional organizations, emphasizing the eight effective teaching practices identified by NCTM (2014) and the eight Standards for Mathematical Practice established in the *Common Core State Standards for Mathematics* (NGA & CCSSO, 2010). The theme-based content and heavily student-centered pedagogy of the Ohio MMR course is new and challenging for most pilot teachers.

The state of Ohio is engaged in ongoing development of instructional materials and PD for this innovative new MMR course. A team of college faculty and high school teachers have worked with Ohio Department of Education (ODE) mathematics coordinator A. Cannelongo and ODE consultant S. Miller to develop course materials and to design PD for teachers piloting the course. The MMR teachers are provided with the curricular scope, sequence, and activities that they are expected to use with their students, and the PD helps them to learn about what they will teach and effective ways to teach it. The in-depth and ongoing PD for Ohio MMR pilot teachers includes face-to-face meetings, online webinars, and individualized meetings with peer mentors and higher education faculty.

The 2020–2021 academic year, the year of data collection for the present study, was the third year of pilot-testing for the MMR course. In 2018–2019, 3 teachers pre-piloted the course. In 2019–2020, 25 additional teachers piloted the course and 25 more piloted the course in 2020–2021. Each year, the MMR teachers provided feedback on the curricular materials for the course and the materials were revised accordingly. The MMR teachers were provided with PD consisting of a summer workshop and follow-up sessions throughout each school year. There were no formal PD sessions in 2018–2019 because the pre-pilot consisted of only three teachers. A 4-day summer workshop occurred in 2019–2020, with a 1-day in-person meeting in January and weekly Zoom meetings in March and April. The 2020–2021 summer meeting was conducted online via Zoom, with follow-up meetings beginning weekly in August. The weekly online meetings shifted to semi-monthly in October and then again to monthly in February. The PD for the MMR course has developed alongside the course materials since the first pilot year. The continuous, ongoing PD has helped the 2020–2021 MMR teachers to teach using the course materials and pedagogy more effectively. Some of these teachers have taught the course and engaged in the PD more than once, reinforcing what they have learned even more.

The MMR pilot teachers have engaged in studying the TAG and MTF as part of their PD workshops through the analysis of written tasks and vignette studies of task implementation. For example, the pilot teachers who attended the summer PD in 2019 and 2020 were instructed to analyze a variety of written instructional tasks using the TAG. Additionally, they were asked to read vignettes of mathematics teaching and identify ways in which teachers used mathematical tasks and student discourse to support student learning. The TAG was introduced as a tool that teachers could use to analyze and reflect on their own instructional practices; this experience positioned the MMR 2020–2021 pilot teachers as ideal research participants for the present study because of the knowledge and experience with the TAG and the MTF that they had developed. Such teachers possessed the potential to select, plan, set up, and implement high cognitive demand tasks and were more familiar with mathematical task terminology than some other high school mathematics teachers in Ohio might have been. Some of the 2020–2021 MMR teachers have taught the pilot course more than once; as such, they have engaged in learning about the MTF and the TAG several times.

Advanced Teacher Capacity

The ATC program provided PD in mathematics and statistics to high school mathematics teachers from 2007–2018. Two PD programs were enclosed under the ATC umbrella: Mathematical Modeling and Spatial Reasoning (Modspar) and Quantifying Uncertainty and Analyzing Numerical Trends (QUANT), which are described briefly in the following paragraphs.

Mathematical Modeling and Spatial Reasoning. Modspar was a yearlong ATC PD program for high school mathematics teachers that consisted of two components: Modeling with Algebra and Modeling with Geometry. Both programs incorporated the use of technology to investigate and solve genuine real-world problems, providing participating teachers with opportunities to set up, engage in, and implement rich mathematical tasks with their students. A central component of Modspar required participants to plan engaging modeling lessons during the summer institutes, teach those lessons during the fall, and report on how they implemented such lessons during the follow-up meetings (Patton College of Education [PCOE], 2019).

As with the Ohio MMR PD, the Modspar program included training and instruction using the TAG and MTF (Stein et al., 2009). Modspar participants read sections of the Stein and colleagues (2009) book and analyzed, planed, and engaged in mathematical tasks individually and collaboratively throughout the summer institutes. Participants were also supported in developing high quality mathematics lessons containing high cognitive demand tasks throughout the yearlong workshops (PCOE, 2019). This training and experience with selecting, planning, setting up, and implementing tasks at high cognitive levels positioned past Modspar participants as ideal research participants for the present study. The Modspar teachers had not only read about the TAG and MTF theoretically but had also applied it to their own instructional practice.

Quantifying Uncertainty and Analyzing Numerical Trends. QUANT was a yearlong ATC PD for high school mathematics teachers, also hosted at Ohio University, focusing on statistics and statistical knowledge for teaching. One of the goals of QUANT was to engage teachers in learning about cognitively demanding mathematical and statistical tasks so that teachers could use such tasks on their own following the PD (Foley et al., 2010). QUANT engaged teachers in reading and using statistics-based adaptations of the Stein and colleagues (2009) TAG and MTF and the *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report* framework (Franklin et al., 2007), a system of (a) formulating questions, (b) collecting data, (c) analyzing data, and (d) interpreting results. Daily activities encouraged QUANT teachers to incorporate the GAISE framework, cognitively demanding statistics tasks, and the integrated use of technology into their instruction.

I also selected QUANT teachers as potential research participants because of the central use of the Stein et al. (2009) TAG and MTF during their PD. Although both are intended for mathematical tasks, they are also applicable for use in statistics (Foley et al., 2010). Though former QUANT teachers were recruited to participate in the present study, none of the 9 participating teachers had been involved in QUANT exclusively (i.e., they were involved in teaching the Ohio MMR pilot course as well).

Data Coding and Analysis

Throughout the present study, I collected and analyzed the following data for each research participant: 1mathematical task, 2 interviews (1 pre-observation and 1 post-observation interview), at least 2 observations (additional lessons were observed for tasks lasting more than 2 days, if possible), and up to 6 samples of student work, depending on how many samples were provided by each teacher. For example, some teachers had fewer than 6 students in their MMR classes or had students submit a single document as a group of 3–4 students. This yielded a total of 9 tasks, 18 interviews, 21 observations, and 36 samples of student work. Data collection and analysis were a continuous, ongoing process (Glesne, 2016; Patton, 2015) as I received and analyzed new data. To best accommodate for teachers' busy schedules, I did not keep all of them at the same stage of the research at the same time; for example, one teacher's post observation interview was scheduled during the same week as another's pre-observation interview and observations. However, data were analyzed as soon as they were collected so that my findings could enhance subsequent observations and interviews (Glesne, 2016; Patton, 2015).

Table 2 provides an overview of the data analysis procedures, including how each type of data was analyzed and the order in which analysis occurred with each teacher's data (vertically from the top of the table to the bottom). The analysis procedures listed in Table 2 occurred once for each research participant, as each submitted 1 main task. First, I analyzed each main task as it appeared in source materials and as it appeared in teachers' lesson plans using Instructional Quality Assessment Academic Rigor Rubric 1: Potential of the Task (IQA Rubric AR1). After transcribing pre-observation interviews, I analyzed them using inductive coding and thematic analysis (Glesne, 2016; Patton, 2015; Privitera & Ahlgrim-Delzell, 2019), an approach that allowed me to identify and organize themes within the data.

Table 2

Data Sources and Analysis Methods

Data Source	Analysis Method
Tasks as they appear in source materials	IQA Rubric AR1
Tasks designed by teachers in their lesson plans	IQA Rubric AR1
Pre-observation interviews	Inductive coding and thematic analysis
Tasks set up by teachers in class	IQA Rubric AR1, EIQA Setup Rubrics 1–2
Tasks implemented by teachers with students (classroom observation and student work samples)	IQA Rubric AR2
Post-observation interviews	Inductive coding and thematic analysis

Note. Data were analyzed in the order they were collected, shown in descending order from top to bottom. Data analysis followed the progression of a single mathematical task through the four task phases: selected from source materials, as planned, as set up, and as implemented. The processes depicted in Table 2 occurred once for each teacher.

Evidence of task implementation included the combination of observations and student work samples, each of which was scored individually using the IQA. After observing instruction and collecting student work samples, I analyzed task setup using IQA Rubric AR1 and EIQA Setup Rubrics 1–2 (CF and MR). Task implementation comprised of two elements: observations and student work samples, each of which I analyzed using Instructional Quality Assessment Academic Rigor Rubric 2: Implementation of the Task (IQA Rubric AR2). Post-observation interviews were analyzed using the same approaches as the pre-observation interviews. The following paragraphs provide explanations for how each data source was analyzed in detail, beginning with interviews because these were the primary data that I used to answer the research questions.

Pre- and Post-Observation Interviews

I used a thematic analysis (Glesne, 2016; Patton, 2015; Privitera & Ahlgrim-Delzell, 2019) approach to coding and analyzing interview data to answer the three research questions. I used an inductive process (Merriam, 2009; Privitera & Ahlgrim-Delzell, 2019) to generate codes and themes because I wanted to keep my analysis focused on the data rather than preconceived notions. Two example interview codes that describe teachers' task selection were "time" and "relevant math content," referring to teachers' tendencies to select tasks based on the amount of time they would take to complete or based on the mathematics content that was addressed.

Coding occurred throughout the data collection process; that is, I analyzed interview data as they were gathered and continued to modify my coding scheme as I collected additional interview data (Merriam, 2009). As I coded interview data, I looked for ways in which codes could be grouped together to form themes. Themes are groups of codes that share common characteristics (Privitera & Ahlgrim-Delzell, 2019), abstractions that are derived from qualitative data (Merriam, 2009). Creation of themes began as an inductive process, followed by a deductive process of assessing themes as new interview data were analyzed. Both the coding of interview data and establishment of themes focused on answering the research questions (Patton, 2015) as I continuously sought ways in which the data might answer them. I developed an initial coding scheme by thoroughly reading, coding, and recoding several interview transcripts and later applied this coding scheme to additional transcripts to determine its effectiveness. I revised the coding scheme and repeated this cycle until no additional codes emerged.

After developing the initial set of codes, I engaged in a second cycle of code work by using codes to create a framework of themes or "relational categories" (Glesne, 2016, p. 200). This stage of thematic analysis required me to reflect on what I had learned and make new connections and insights (Glesne, 2016). This process resulted in the development of themes, words or phrases that describe codes (Privitera & Ahlgrim-Delzell, 2019). After identifying themes that emerged within single cases (i.e., single teachers), I used cross-case thematic analysis, the process of "identifying patterns across cases and assigning them a name to describe a pattern" (p. 743), to compare codes and themes across multiple teachers.

I stored and managed interview data using NVivo, a software package that allows users to transcribe, read, and code interview data. As I coded interview transcripts, I kept a running list of major codes and sub-codes, including descriptions of each, in a digital codebook (Glesne, 2016; Patton, 2015) using Microsoft Excel. This enabled me keep to the original meaning for each code and determine whether codes could be combined, split, modified, or removed (Glesne, 2016).

Mathematical Tasks Across the Four Task Phases

To determine the cognitive demands of tasks as they appeared in source materials, as planned, as set up, and as implemented, I collaborated with another trained graduate student at Ohio University to use the IQA instrument. Both analysts applied IQA Rubric AR1: Potential of the Task to analyze the cognitive demand of submitted tasks as they appeared in source materials, as they appeared in teachers' plans, and as they were set up for students during lesson observations. When analyzing tasks as they appeared in source materials, both analysts assigned a numerical IQA score (0, 1, 2, 3, or 4) to the digital version of the source tasks submitted by participating teachers, reaching a consensus for each task through dialogue if we initially disagreed. This score, referred to henceforth as a *Phase 1: Selection* IQA score, was given for each teacher's task as it appeared in source materials. Phase 1: Selection indicates the first phase in the four-phase task model provided in Figure 1 on page 50, tasks as they appear in source materials, whereas IQA in the score name indicates that the selected task was scored by the researchers using the Instructional Quality Assessment (as opposed to being scored by participating teachers using the Task Analysis Guide).

To analyze tasks as planned, both analysts assigned a similar score to each task as it appeared in teachers' lesson plans, again reaching a consensus for each score after discussion. Not all the participating teachers wrote official lesson plans, but this phase also included modifications that teachers made to their tasks (if any) when planning their instruction. These scores are referred to as *Phase 2: Planning IQA* scores, signifying task scores for the second phase of Figure 1, tasks as planned by teachers. Both analysts also used IQA Rubric AR1 to assign unanimous scores for task setup; such scores are referred to as *Phase 3: Setup IQA* scores. Analysts assigned Phase 1: Selection, Phase 2: Planning, and Phase 3: Setup scores according to the level of thinking required of students to provide a "complete and thorough response that satisfies the stated demands of the task" (Boston & Wilhelm, 2017, p. 841), determined and agreed upon by both trained IQA raters during the analysis. '

Task setup received two additional scores, a *Setup 1: CF* score and a *Setup 2: MR* score, representing scores using EIQA Setup Rubric 1: CF and EIQA Setup Rubric 2: MR, respectively (these rubrics were described in Chapter 2). Consistent with the standard IQA, scores ranged from 0–4 and aligned closely to the Task Analysis Guide. The Phase 3: Setup IQA score for each task describes the cognitive demand of the task itself during the setup phase; that is, this score is meant to reflect whether the cognitive demand of the task from the previous phase is maintained, declines, or inclines. However, the two EIQA scores address particular components of whole-class discussions that occur as part of task setup (if such discussions happen). Therefore, the two EIQA scores capture a subset of the instruction captured by the Phase 3: Setup IQA score for each task.

Recall that task implementation was defined in the present study as the verbal and written communication made between the teacher and students while solving the task and the work that is done to complete the task (Stein et al., 1996, 2009). Therefore, evidence of task implementation included the observational field notes and student work samples pertaining to each task. Triangulation between observational data and student work samples provided a thorough analysis of tasks as they unfolded during instruction because the strengths of one approach supported the weaknesses of the other; the observations provided evidence of what students said and did during instruction whereas the student work provided evidence of students' thinking and the work they actually accomplished (Boston, 2012; Glesne, 2016). Both of these data sources were analyzed using IQA Rubric AR2, Implementation of the Task. As with Rubric AR1, scores of either 0, 1, 2, 3, or 4 were assigned to each observation and set of student work pertaining to a single task.

Task implementation scores should be "holistic, reflecting the highest level of engagement of the majority of students during individual or small-group work on the task" (Boston & Wilhelm, 2017, p. 841); therefore, I first assigned a single, holistic *Implementation: Student Work* score based on the majority of the student work samples pertaining to each task. I also assigned a *Phase 4: Implementation IQA* score to each set of observations related to a single task. That is, I analyzed each teacher's observations holistically and provided a single IQA score based on the mathematical work and thinking that the majority of students were doing for most of the class time. I chose not to combine Implementation: Student Work scores and Phase 4: Implementation scores by averaging them because each score represents a qualitative category with unique qualitative attributes; each score categorizes a different type of data, one pertaining to documents and the other pertaining to observations. As such, combining scores did not seem logical. In summary, each task was assigned 7 IQA scores: 1 for Phase 1: Selection, 1 for Phase 2: Planning, 3 for Phase 3: Setup (2 of which are EIQA scores), and 2 for Phase 4: Implementation (1 for observations, 1 for student work).

The third research question concerns mismatches in task analysis between the researchers and participating teachers. A mismatch, for the purpose of the present study, occurs when the researchers and a teacher disagree on the level of cognitive demand of a task at any of the four task phases. However, the IQA and TAG are not identical in terms of the level of cognitive demand evident in each classification; though IQA scores of 1 and 2 align with the TAG categories of *memorization* and *procedures without connections* tasks, an IQA score of 3 may be assigned for *procedures with connections* tasks and *doing mathematics* tasks. Similarly, the IQA score of 4 may be assigned for both types of tasks if they provide explicit opportunities for students to provide evidence of their mathematical thinking and reasoning.

Because of the slight differences in the structuring of the IQA and the TAG, a mismatch is therefore defined as an instance where the researchers and a teacher disagree on whether the cognitive demand of a task is high at a specific task phase; according to the TAG, *memorization* and *procedures without connections* tasks are low in cognitive demand, whereas *procedures with connections* and *doing mathematics* tasks are high in cognitive demand. This suggests that IQA scores of 1–2 categorize low cognitive demand tasks and scores of 3–4 categorize high cognitive demand tasks. Mismatches in task analysis are then evident if either (a) the researchers provide an IQA score of 1 or 2 for a task phase and a teacher classifies the same phase as *procedures with connections* or *doing mathematics* or (b) the researchers provide an IQA score of 3 or 4 for a task phase and a teacher classifies the same phase as either *memorization* or *procedures without connections*. Tasks that are scored as 0 using the IQA are analogous to *nonmathematical activity* tasks according to the TAG and are both considered to be low in cognitive demand as well.

Researcher as Instrument and Positionality

"Reflexivity" (Merriam, 2009, p. 219), otherwise known as positionality, refers to reflecting on subjectivities and how they influence the researcher's position and the research itself, a critical reflection of the self as researcher. Because the researcher is the instrument of qualitative research (Brinkman & Kvale, 2015), credibility and is established by developing integrity as a researcher (Patton, 2015) and explicitly describing how a researcher's beliefs and values influence the study (Merriam, 2009). As the researcher in this study, I acknowledge that my experiences as a student and as a teacher of mathematics influence my research interest in mathematical tasks and my pedagogical beliefs about cognitive demand. My high school and undergraduate courses in mathematics had been primarily taught through lecture, with emphasis on the ability to execute mathematical procedures fluently. Though I was able to make sense of concepts on my own, I did not understand why other students had difficulty with this approach. As an undergraduate student in mathematics education and as a graduate teaching assistant, I taught the same way I was taught. Even at the master's level, all my mathematics courses were taught by lecture, further enforcing the notion that lecturing and attention only to procedural fluency (with little to no emphasis on conceptual understanding) were the ways of teaching mathematics. Teaching as telling (Philipp, 2007; Romagnano, 1994) was the norm for me.

It was not until my coursework as a doctoral student in mathematics education and my experience with cognitively demanding tasks that my beliefs about mathematics teaching changed. Through my reading of the literature involving mathematical tasks, engagement in PD with mathematics teachers using the TAG and MTF, and involvement at state and national conferences for mathematics teachers, I have developed the strong belief that students should be creators and doers of mathematics, whereas teachers should serve as facilitators who teach for both procedural fluency and conceptual understanding through the regular use of high cognitive demand tasks. I mention my educational background and beliefs because they influence how I perceived and analyzed mathematics instruction and mathematical tasks during this research. My personal experience in high school mathematics classrooms and reading of the literature influence my perception that high school mathematics instruction across the state of Ohio and the United States was frequently taught through lecture and focused primarily on the memorization of facts and the correct execution of routine procedures (Stigler & Hiebert, 1999; Stein et al., 2009).

Reflexivity as a researcher requires me to reflect on my beliefs and preconceived notions throughout this research process. Rather than assuming that my participants held beliefs evident in the literature or beliefs that may be typical, I have grounded my analysis on what my participants convey to me through interviews, as reported in Chapter

4. The purpose of this study was not to determine whether the perspectives held by research participants were similar to those evident in the literature, nor was the purpose to assess such perspectives. Rather, the aim of this study was to learn from my participants and to develop a deeper understanding of how their perspectives influence their decisions when selecting, planning, setting up, and implementing mathematical tasks. The nature of the present study was exploratory, as opposed to evaluative, and my role as the researcher was to illicit and report what my participants communicated. However, my beliefs and values motivated this research and contributed to my desire to learn more about the complex interactions between mathematics teachers and the tasks they use during their instruction.

Qualitative researchers not only introduce their own biases and preconceptions, but also benefit research through their expertise, knowledge, and experience. Interviewing, in particular, is a skill that can be learned and developed over time by reading qualitative research reports and texts, conducting qualitative research, and practice (Glesne, 2016; Patton, 2015). Patton (2015) asserts that qualitative researchers are *connoisseurs*, stating that "the researcher as connoisseur or expert uses qualitative methods to study a program or organization but does so from a particular perspective, drawing heavily on his or her own judgments" (p. 210). My teaching and research experience, mathematics knowledge, and specifically, my work involving mathematical tasks, has equipped me with the qualitative research skills necessary to conduct the present study. I have taught mathematics at the middle school, high school, and postsecondary levels, attended and co-led teacher PD, given conference presentations on mathematical tasks and facilitated teachers' analyses using the TAG, and conducted qualitative interview research throughout my career as a mathematics educator. My work with high school mathematics teachers through task-focused conference presentations has enhanced my ability to analyze mathematical tasks and communicate effectively with teachers regarding such topics.

Credibility, Transferability, and Trustworthiness

Though validity and reliability are terms that are generally applied to quantitative research, some authors of qualitative research methods texts advocate for their use as well. According to Merriam (2009), "internal validity deals with the question of how research findings match reality. How congruent are the findings with reality? Do the findings capture what is really there?" (p. 213). However, the notion of "reality" varies with qualitative research; Merriam writes that "one of the assumptions underlying qualitative research is that reality is holistic, multidimensional, and ever-changing" (p. 213). Credibility, "the correspondence between research and the real world" (Wolcott, 2005, p. 160), is often the term employed by qualitative researchers instead. Both terms communicate the necessary connection between the research and what is actually being studied, however, credibility acknowledges that a single "valid" result does not necessarily exist. Qualitative researchers instead seek the "truths" (Rubin & Rubin, 2012, p. 10) based on the lived realities of their research participants. Merriam (2009) offers the following methods to enhance the credibility of a study: triangulation, member checks, adequate engagement in data collection, reflexivity, negative case analysis, and peer review, each of which is discussed in the following paragraphs. This subsection concludes by addressing transferability and trustworthiness.

Triangulation is the use of multiple methods and multiple data sources to enhance the credibility of qualitative research (Glesne, 2016; Merriam, 2009; Patton, 2015). The present study incorporated the use of multiple methods of data collection to support data analysis through the collection of classroom artifacts (student work samples), observations, and teacher interviews. Each data collection method provided checks to the others and provided a nuanced understanding of each mathematical task as it unfolds during instruction. I compared interviews to written field notes during observations, providing clarification of the events that I observed and further insights into teachers' task selection, planning, and setup. In this way, each has served to deepen the interpretation and understanding of the others (Glesne, 2016).

Triangulation between data collection methods and data sources also enhances credibility by using the strengths of one data source to compensate for the weaknesses of the others (Patton, 2015). For example, observations of instruction cannot possibly capture teachers' mental processes and rationales; however, I have gained this understanding through post-observation interviews and by connecting teachers' responses to what I had observed. Student work samples alone did not capture the cognitive demands of tasks as implemented but were combined with observational data to provide a thorough analysis of the work that students produced in solving each task (Boston, 2012).

Member checks, or "respondent validation" (Merriam, 2009, p. 217) supports credibility in qualitative research by verifying the accuracy of the researcher's interpretation. This strategy helps to confirm whether the data collected and analyzed represent participants' truths and realities. Member checks involve sharing interview transcripts or even preliminary analyses back to research participants and asking them to verify that what has been described by the researcher is accurate (Glesne, 2016; Merriam, 2009). Sharing interview transcripts and observational field notes can also enhance the trustworthiness and rigor of a study (Glesne, 2016). Throughout the present study, I conducted member checks of interview transcripts with each of my research participants to ensure that the transcripts accurately reflected teachers' thoughts, decisions, and actions. After transcribing each interview, I shared it with the appropriate teacher and asked them to verify the accuracy of what was transcribed.

Negative or discrepant case analysis (Glesne, 2016) is the purposeful search for "data that might disconfirm or challenge your expectations or emerging findings" (Merriam, 2009, p. 219). This technique establishes credibility by seeking to eliminate the possibility of alternative explanations (Patton, 2015) to "increase confidence in the original, principal explanation you generated" (Merriam, 2009, p. 219). By eliminating other possibilities, qualitative researchers can thereby support their analyses and explanations related to the phenomena of interest. I have done this in this study by asking questions to participating teachers during post-observation interviews. After observing the implementation of a mathematical task in the classroom, I likely had my own interpretation(s) for how the task unfolded and why; however, I was able to eliminate some possibilities by asking interviewees for their interpretation, seeking further clarification for what I might have missed. I explored alternative explanations for task implementation by thoroughly examining student work and interview transcripts, searching for answers to questions that I still had after observing.

The last of Merriam's (2009) suggestions for enhancing credibility is the use of peer review or peer examination. This involves obtaining external input (Glesne, 2016) and asking others whether research findings are plausible within a given data set (Merriam, 2009). As a graduate student working to complete a dissertation, I am fortunate to have been able to present updates on data collection and analysis to the faculty on my dissertation committee. I was provided feedback from various perspectives, including that from the fellow graduate student who observed and analyze instruction using the IQA for reliability purposes; throughout the data collection and analysis process, I consulted with this trained rater to evaluate the accuracy of my analyses and the plausibility of the findings that emerged.

Unlike credibility, or internal validity, external validity requires that the findings of a study be applicable to other situations (Merriam, 2009). However, qualitative research enhances the understanding of small samples (Patton, 2015) and does not necessarily attempt to generalize results from a sample to a larger population. Most qualitative research relies on purposeful, rather than random, sampling. Some believe that is better to consider "transferability" (Lincoln & Guba, 1985, as cited in Merriam, 2009, p. 224), the idea that consumers of qualitative research should be left to determine whether the research findings can be applied to other situations. However, it is the qualitative researcher's responsibility to provide enough data to make transferability possible (Lincoln & Guba, 1985, as cited in Merriam, 2009). This has been achieved through the “thick, rich description” (Glesne, 2016, p. 153) that is the essence of qualitative research and transparency in reporting research decisions, processes, and analyses.

Transferability is also enhanced through the use of an audit trail, a method introduced by Lincoln and Guba (1985) that describes how data were collected, how themes were derived, and how decisions were made throughout the research (Merriam, 2009). Throughout data collection, analysis, and interpretation, I kept a research journal in which I explained my decision-making, reflections, and questions and while collecting and analyzing data. I have been transparent in all my activities, procedures, and decisions as I recorded all this information in my research journal and presented it in this dissertation.

Glesne (2016) suggests that qualitative research addresses trustworthiness rather than validity and cite Lincoln and Guba’s (1985) use of credibility, transferability, dependability, and confirmability as constructs to assess trustworthiness. “These constructs parallel ones used in quantitative research (internal validity, external validity, reliability, and generalizability) but employ different strategies for different ends” (Glesne, 2016, p. 152). Trustworthiness in this study has been developed through my professional connections to some of the potential research participants prior to conducting the research.

Trustworthiness has also been established through communication and collaboration with research participants throughout the study. I was flexible with teachers’ schedules when setting up observations and interviews and I conveyed the message that this study serves to enhance high school mathematics teachers' ability to teach cognitively demanding tasks and reflect on their instruction. Moreover, Merriam (2009) asserts that trustworthiness in qualitative research involves conducting the research in a trustworthy and ethical way. I have conducted an ethical study by adhering to the IRB requirements to conduct research and the IRB protocol I have created; for example, the names and identifies of participating teachers and their students have remained anonymous through the use of pseudonyms and removal of names on student work samples. I also reminded each participant that their involvement in this research study was voluntary and that they could withdraw from the study at any time without penalty.

Chapter Summary

The present study is situated within the context of teacher professional development, including the Ohio MMR pilot course and two ATC PD programs, QUANT and Modspar. Through the selection of participants from such contexts, I aimed to recruit participants who had experienced the MTF and TAG in PD and were familiar with the terminology related to QUASAR research. Qualitative interviews served as the primary source of data to answer the research questions, providing evidence teachers' perspectives regarding the use of tasks at each phase and factors that they attributed to changes in cognitive demand between each phase. Interview data were augmented through the collection and analysis of mathematical tasks at each phase, student work samples, and observations, which were analyzed using the IQA and EIQA rubrics.

Finally, I conducted both a rigorous and ethical study by establishing credibility, transferability, and trustworthiness.

Chapter 4: Findings

In this chapter, I present the findings from the analysis of interviews, remote classroom observations, and mathematical tasks across the four phases of selecting, planning, setting up, and implementing. The first section describing the individual cases includes 9 parts, one for each participating teacher. The findings for each teacher include:

- the IQA analysis of the teacher's task across the four task phases,
- the teacher's analysis of the task at each of the four phases using the TAG, and
- the teacher's reasons for selecting, planning, setting up, and implementing tasks, including both general reasons and those that pertain to their chosen task for the present study.

Next, I present the findings from the thematic analysis across cases. This begins with a description of the 18 themes identified through the coding of interview data. The 18 themes are sorted by task phase, with 5 themes for task selection, 5 for planning, 3 for setup, and 5 for implementation. After discussing the themes derived from interviews, I describe the cross-case analysis of mathematical tasks, which is divided into three parts: (a) trends in IQA scores across task phases, (b) trends in teachers' TAG classifications across task phases, and (c) mismatches between IQA scores and TAG classifications.

The Individual Cases

In the following sections, I describe the 9 participating teachers and the instructional tasks they chose for me to observe. I begin with a description of each teacher's background, including their teaching experience and role in the MMR pilot course and ATC programs. To provide context for each teacher's interview responses, I present the teacher's chosen task and explain the changes that were made to the task during planning, setup, and implementation. Next, I present the IQA analysis of a teacher's task at each phase, followed by the teacher's analysis using the TAG. I then explain each teacher's reasons for (a) selecting tasks from source materials, (b) planning and adjusting their tasks accordingly, (c) setting up tasks for student engagement, and (d) implementing tasks with their students. Finally, I provide reasons that each teacher attributes to the cognitive demand of each task and reasons that they attribute to the change in cognitive demand across the four task phases.

A pseudonym is used for each teacher and direct quotes from teachers are set in quotation marks. Throughout the following descriptions of teachers' tasks, I often refer to *Contexts* and *Themes*, terms that are used to categorize the MMR curriculum. An MMR *Context* refers to a set of tasks or activities centered around a given real-world context (e.g., ramps) and is relatable to a *chapter* in a typical mathematics textbook. Analogously, a *Theme* in MMR refers to a strand of mathematics content, comparable to a typical *unit* (e.g., Number and Quantity). Themes consist of approximately 5 Contexts, and each Context is typically meant to span several 45-min lessons.

Tables 3 and 4 provide a synopsis for the following sections of Chapter 4, displaying the IQA and TAG scores for each teacher's task, respectively. As described in Chapter 3, my co-observer and I assigned the following scores for each teacher's task: (a) an IQA score for each of the four phases of teachers' task use (Phase 1: Selection–4 scores), (b) expanded IQA (EIQA) scores for the discussion of contextual features (Setup

1: CF) and mathematical relationships (Setup 2: MR) evident in task setup, and (c) an IQA score for each set of student work (Implementation: Student Work).

Table 3

Instructional Quality Assessment Scores for Teachers' Tasks

Teacher	Phase 1	Phase 2	Phase 3	Setup 1: CF	Setup 2: MR	Phase 4	Student Work
Adam	4	4	NR	NR	NR	2	NR
Beth	3	3	3	NS	NS	3	2
Cathy	3	3	3	3	2	2	3
Debbie	4	4	4	3	3	4	4
Ethan	4	4	4	NR	2	3	NR
Fred	4	4	4	2	2	4	3
Gwen	3	3	4	3	3	2	2
Henry	3	3	3	NA	NS	4	3
Isabel	4	4	NR	NR	NR	NR	NR

Note. Phase 1, 2, 3, and 4 scores refer to task selection, planning, setup, and implementation, respectively. "Implementation: Student Work" is condensed to "Student Work" in the last column so that the table fits within the margins of the page.

Table 4

Task Analysis Guide Classifications for Teachers' Tasks

Teacher	Phase 1	Phase 2	Phase 3	Phase 4
Adam	DM	PwC	PwC	PwC
Beth	DM	DM	DM	DM
Cathy	DM	DM	DM	DM
Debbie	PwC	PwC	DM	PwC
Ethan	PwC	PwC	DM	PwC
Fred	PwC	PwC	PwC	PwC
Gwen	PwC	PwC	PwC	PwC
Henry	DM	DM	DM	DM
Isabel	PwC	DM	DM	DM

Note. Phase 1, 2, 3, and 4 scores refer to task selection, planning, setup, and implementation, respectively. The TAG classifications presented in Table 4 were designated by the research participants rather than the researcher. Each teacher classified their own task based on the level of cognitive demand that they perceived their task to possess. The TAG classifications in the following tables were also assigned in this way.

In Tables 3 and 4, classifications of "NR" (no rating) were assigned for IQA task phases that could not be scored because either (a) they were not observed, (b) were qualitatively unique and not sensible to score using the IQA rubrics, or (c) in the case of students' work, lacked the minimum 4 samples of student work to score reliably. The rating "NC" which represents "no classification" was assigned for Setup 1: CF, Setup 2: MR, and Implementation: Student Work when assessed by teachers using the TAG because the TAG applies specifically to task Phases 1–4. Additionally, scores of "NS" and "NA" (in place of N/A) were used as intended by the EIQA rubrics, presented in Appendix B. The TAG classifications of *procedures with connections* and *doing mathematics* are abbreviated as "PwC" and "DM," respectively, to condense the width of Table 4. These labels and abbreviations are used throughout the remainder of this chapter.

Adam

Adam has taught school mathematics for 14 years, all of which have been at the same high school. The 2020–2021 academic year was his first year teaching the MMR course. Between two sections of the course, Adam has a total of 25 students. All of them are seniors who chose not to take precalculus, statistics, or the "college math class," offered at their school which is "really just a review of Algebra 1 and Algebra 2 which all of them [Adam's students] already had." Due to the COVID-19 pandemic, Adam's district had him teaching from home through Zoom with "about 90%" of his students connecting online from his classroom and the remaining students connecting remotely from home. More of his students were previously learning remotely, but some "found that they needed to transition back to face-to-face" because they were "not doing so good" and did not "feel like they were being held accountable… for doing work outside of class." My co-investigator and I observed Adam teach a section of 9 students on consecutive days via Zoom. The IQA and TAG classifications for Adam's task are presented in Table 5 to provide context for the following two sections.

Table 5

IQA Scores and TAG Classifications for Adam's Task

Score Level	IQA Score	TAG Classification	Mismatch
Phase 1: Selection	4	DM	No
Phase 2: Planning	4	PwC	No
Phase 3: Setup	NR	PwC	No
Setup 1: CF	NR	NS	No
Setup 2: MR	NR	NS	No
Phase 4: Implementation	2	DM	Yes
Implementation: Student Work	NR	NS	No

Recall that a *mismatch* in task analysis refers to a difference in the classification of a task by the researchers using the IQA and the teacher using the TAG at any one of the four task phases. Specifically, a mismatch occurs when a task that is classified as high-level (scores of 3–4 on the IQA or *procedures with connections* and *doing mathematics* in the TAG) by one party is classified as low-level (scores of 1–2 on the IQA or *memorization* and *procedures without connections*) by the other. As I discussed in

Chapters 2 and 3, IQA scores of 3 and 4 do not directly correspond to the TAG categories of *procedures with connections* and *doing mathematics*; that is, an IQA score of 3 may be suitable for a *procedures with connections* or a *doing mathematics* task. Because the defining characteristic between IQA scores of 3 and 4 is explicit evidence of students' thinking and reasoning, *mismatches* are defined as differences between low cognitive demand tasks and high cognitive demand tasks.

Analysis of Adam's Task. The task that Adam chose for the observations was titled "Remodeling Our Classroom," a task that was provided to MMR teachers in the course materials. The context for the task is that students are to submit a proposal to redesign their classroom space, including carpet, a new coat of paint, and a gas fireplace. Some parameters are provided in the task statement, such as the cost of paint, carpet, the gas line, and the gas fireplace. However, students are given the freedom to incorporate additional elements of their choice. The task consists of three parts: (a) making an initial estimate for the total cost of project materials, (b) submitting and presenting a budget for approval, and (c) creating a diagram of the remodeled classroom to accompany the budget proposal. Students are also provided with specifications for how the gas line should run through the classroom and are expected to use 2 coats of paint, assuming that 1 gal of paint covers 500 square feet.

The Phase 1: Selection IQA score for this task as it appears in curricular materials is a 4 because it has the potential to engage students in exploring mathematical relationships and encourages the use of complex, non-algorithmic thinking. Though students are required to include the cost of paint, carpet, the gas line, and the gas fireplace, they are not told how to find these costs for the given room they are working with. The task is open-ended and invites students to develop their own solution pathways using their prior knowledge of surface area and dimensional analysis. There is not a previously worked-out example for how to do the calculations or what calculations should be made in the process. Moreover, the task explicitly calls for evidence of students' thinking and reasoning in their proposals through the following statement: "Be sure to thoroughly document and explain your work. Use pictures, formulas, units, tables, etc. to make the cost calculations as easy for others to understand as possible." The task also states that students should be prepared to present their proposals to the school board or administration for approval, adding a verbal requirement for evidence of students' thinking and reasoning.

Because Adam did not make any changes to the task when planning, the Phase 2: Planning IQA score for this task is also a 4. Adam did not modify the task in any way and used the same handout that was provided with the MMR materials.

After consulting with an IQA expert, I have chosen not to score Adam's task setup (Phase 3: Setup) because the setup of the task occurred on a day that I did not observe. During the 2 days that I observed I only saw the implementation of the task, which lasted both days. Instead of assigning a numerical IQA score, I have assigned the code NR (no rating given) for the Phase 3: Setup IQA score and for Setup 1: CF and Setup 2: MR (NS, representing *no score*, is reserved for different circumstances within the EIQA rubrics). However, I asked Adam interview questions to develop an understanding of his perspective of the task as set up; a synopsis of the task at set up is provided in the following paragraphs.

To help introduce the task and the handout, Adam provided his students with the task rubric, provided with the MMR course materials, and communicated his expectations for students as they worked in groups. He then shared a video on "how to do scale drawings" to guide students' work on that part of the task. After watching the video, Adam led his students in a discussion of what they noticed and what they wondered, a practice that is emphasized in the MMR course as an avenue to engage students in mathematical discourse. Next followed an estimation activity where students individually predicted one surface area that they knew would be too high for the classroom walls, one that they knew would be too low, and one "educated guess in the middle."

The students then split into groups and measured the dimensions of the classroom. Adam noted how some students' estimations were "so far all over the place" whereas others' were "within a hundred… square feet of what the classroom actually was." After a brief discussion of students' findings, Adam spent some time discussing some of the real-world aspects of the task: he acknowledged that both a fireplace and carpet were unrealistic to include in a classroom, but that students would apply what they learned through the task later in life if they needed to install a fireplace or carpet in their homes, for example. The class also discussed the dropped ceiling in the classroom because of how it would interfere with the gas line that students would design as part of the task. According to Adam, the discussion helped his students to visualize where the piping for the fireplace would go.

Then, Adam and his students determined the length of the gas line together as a class. Using his iPad, Adam drew a three-dimensional sketch of the classroom indicating where the gas line would go based on the specifications given in the task handout. However, he did not provide his students with the measurements, they had to get up out of their seats and measure the pipeline themselves. Adam drew a series of segments to indicate where the gas line would be positioned in the room and asked students to provide the dimensions of the piping until they had determined its entire length. Finally, Adam gave students "free rein" to do their calculations, identify other items they wanted to include, and develop their proposals. In doing so, he verbally set the expectation that students were to justify their work with evidence. He said,

> If they [students] want to remove a chalkboard off the wall, okay, remove the chalkboard off the wall… but you better make sure that you can go somewhere and find the labor and the cost that's going to represent to do that.

Adam made it clear that students were required to provide documentation for the costs involved in their proposals and that they needed to include all the calculations they made.

For the purpose of this study, task implementation includes evidence from (a) observations of teachers' instruction and (b) evidence of students' written work on a task. Each element is assigned a separate score, labeled as Phase 4: Implementation IQA and Implementation: Student Work scores, respectively. For Adam, the former was assigned a score of 2 because there was little ambiguity about what was expected of students mathematically and there was little evidence of students' work and thinking processes throughout the 2 lessons. Students worked in groups (one group of 4 students and another group of 5) to perform their calculations for the cost of paint, carpet, and the gas line, but students did not explicitly make connections to mathematical concepts underlying the procedures they used. They divided the work of the task into parts with each student completing their own section and focused primarily on numerical calculations. Some aspects of the task were decided by students without prompting and direction from the teacher, but they were not mathematical in nature. For example, one group decided that they would need to purchase a toolkit because it would include all the tools that they might need to remove the chalkboards from the walls themselves. They decided that they would paint under, rather than around, the chalkboards to simplify the calculations for the amount of paint they would need.

Rather than directly showing students what calculations to make and how to do them correctly, Adam acted as a facilitator as he moved periodically from one online breakout room to the other. He monitored students as they worked on their written proposals and rarely stopped to intervene with a question or comment. Because students' work was not visible from our perspectives as Adam, my co-observer, and I observed through the Zoom meeting, Adam periodically asked what students were working on.

One example of a typical conversation is the following: Adam asked one group how they figured out the total cost for the paint to cover the classroom walls. When one student answered that she was working on that part, Adam asked her how she figured out the total cost for the paint; she replied that she multiplied the height of the classroom by its width and then multiplied that number by 4. In response, Adam asked what she was going to do about obstacles in the classroom, such as the chalkboards and the door, and attempted to direct her toward subtracting their surface area from the total. However, the group decided to manually remove the chalkboards and paint behind them. Adam had no problem with this choice, if the group included the associated costs in their proposal. He then allowed them to continue on with their work without asking the group to explain their mathematical reasoning.

Adam facilitated groupwork in this way throughout the majority of the t class meetings, remaining predominantly silent and occasionally inquiring about students' plans and progress. Some students discussed what additional items they would include in the proposal, if any, and asked one another to check over their calculations. At the beginning of class on the second day, Adam reiterated his expectations for the scale drawing component of the task and provided rough sketches of two- and three-dimensional layouts of the classroom to help students visualize what he was looking for. Due to scheduling conflicts with other teachers, my co-observer and I did not observe students' presentations that were set for the following week.

To assign an Implementation: Student Work score for a task, a minimum of 4 samples of student work is necessary according to the literature and the recommendation of the IQA expert whom I consulted throughout the present study. The 9 students in Adam's MMR class completed the task in groups, resulting in only 2 samples of work. Without a minimum of 4 samples, the Implementation: Student Work score assigned for Adam's task is NR. However, a qualitative analysis of the 2 samples shows that each group engaged in a different level of mathematical work and thinking. One group's work, Sample A, includes a two-dimensional (digital) layout of the classroom that is clearly labeled with a door, chalkboards, desks, and other objects. However, the costs of carpet and paint are given without calculations and without any written justification. Measurements lack units and several two-dimensional quantities are given with one-dimensional units (e.g., the surface area of the door is given in feet rather than square feet). In general, Sample A lacked evidence of students' thinking and required some interpretation to understand.

Sample B, though lacking a scale-drawing of the classroom (perhaps the group had not finished that part yet), included more sophisticated mathematical work and reasoning to justify it. For example, the calculations for the surface area of the chalkboards, door, and walls were presented clearly with appropriate units and correct conversions from one unit to another. This group provided the total surface area of the room (864 square feet) and indicated that they needed to cover 679.74 square feet with paint, the result of subtracting the surface area of the chalkboards, the door, and the lockers lining one of the walls (120 square feet, 29.7 square feet, and 34.56 square feet, respectively) from the total. Sample B included reasoning for students' calculations to support the amount of paint needed to cover the walls, though they forgot to double the amount of paint for the second coat in their calculations. The students in this group provided a brief written explanation to summarize their work and included documentation of the additional items they would include in their proposal.

Adam's Analysis of the Task. When asked to classify the MMR-provided version of the "Remodeling Our Classroom" task, Adam stated that it was a at the level of *doing mathematics*. He specifically pointed out that the task requires "a lot of self-monitoring" and "self-regulations" as students must determine what to include in their proposals, where and how to find the information, and how to present the appropriate reasoning and documentation to support their choices. The mathematics to determine the amount of the paint on the walls, according to Adam, was not as straight-forward as multiplying the length of a wall by its height and then multiplying by 4; students needed to subtract out the surface area of chalkboards, cabinets, and other objects in the room but also for additional items they might include in their proposals, such as posters, which may vary from group to group. Students also must "analyze the task and accurately examine the task constraints" throughout the task. Adam provided the example that students cannot simply order desks without making the appropriate measurements to determine whether they will fit in the classroom space as they are arranged.

Adam did not make adjustments to the original version of the task and its cognitive demand remained the same in Phase 2: Planning. After setting up the task, Adam expressed that the cognitive demand made a "transition from *doing mathematics*… to a higher… *procedures with connections* task," communicating a misconception that *procedures with connections* was at a higher level than *doing mathematics*. I chose not to explain the ordering of the TAG levels of cognitive demand because I did not want to influence his response. Adam explained that his students would execute the procedures of researching, gathering data, and representing data using spreadsheets. According to Adam, having students use spreadsheets would also support connections to data because the spreadsheet formulas would reflect associated item costs and quantities. Subtracting out the surface area of the chalkboards, bookshelves, and lockers in the classroom would force students to connect ideas about surface area together. Adam reflected that, in general, the cognitive demand of a task "always increases" from its representation in source materials because "in the beginning, it's just a handout with instructions on it." As students begin to engage with a task, their increased work and thinking drives the cognitive demand upward in his view.

Adam claimed that the implementation of his task was still "at the high level in terms of making connections." After working on the task for several days, Adam was pleased with the connections he felt that his students were making in terms of representing their data in various ways (scale drawings, spreadsheet formulas, numerical costs), calculating and reflecting on the reasonability of unit conversions, and making real-world connections (e.g., the reality that including a fireplace and carpet would not make sense in a school classroom). A key feature of task implementation that set it "out of the lower-level categories and definitely a higher-level" was that Adam did not manipulate the direction that students took when solving the task and allowed them to make their own decisions. Adam attributes the continued high cognitive demand of task implementation to the conversations that students had with each other as they worked collaboratively toward completing the task and critiquing their own and each other's work. Though working independently and dividing up the work of the task, students organized and connected the different pieces when preparing their proposals.

Phase 1: Task Selection. The time it takes students to complete a task is Adam's first consideration when selecting tasks to use for his classes. "Right now, the number one thing is time. How much time is it going to take to do it?" In a typical year, the MMR course would be more tightly structured, and teachers would be required to progress from one task (or *Context*, using MMR language) to the next, following the general pacing guide provided with the course materials. However, due to the COVID-19 pandemic, the MMR teachers were given the "freedom to pick through the lessons" and choose which tasks they wanted to use; in a typical year, the Ohio Department of Education (ODE) would have conducted research pertaining to the course and required MMR teachers to follow the scope and sequence of the course strictly. Because the COVID-19 pandemic limited the ability to collect data and conduct research, MMR teachers were allowed such freedom to pick and choose lessons. According to Adam, "At the pace that I was at, I would still be in Theme 0 (the first MMR Theme for the course) and I would be weeks behind because it was typically set up to last 3 weeks… we were already in school almost 6 weeks." Adam's adjusted teaching schedule for MMR was 4 days per week for 51 minutes, 1 day short of the typical 5-day school week. Therefore, time dictated what Adam felt he was able to do and how long he had to spend on a given task.

Adam's remote teaching environment also influences his selection of tasks, specifically for the MMR course. When selecting a task, he wonders, "Are the kids going to be engaged? Will they make connections? And so, I… go in a little apprehensive, thinking, will this work?" Adam acknowledged that his students might have engaged with instructional tasks differently based on whether they were learning remotely or face-to-face. "I have kids… in class [MMR] that join virtually and those that are in my classroom at the same time… can I make this fit for both groups?" As several participating teachers suggested throughout this study, some tasks might not be suitable for students who attend school virtually or for those in hybrid environments where students are split between face-to-face and remote groups. Though Adam taught remotely from home, most of his students attended class at school unless they were quarantined; therefore, the "Remodeling Our Classroom" task was suitable for Adam's students.

Feedback from other teachers piloting the MMR course helps Adam to determine which tasks to use and which to avoid. Throughout the 2020–2021 academic year, the cohort of MMR pilot teachers met biweekly to discuss the lessons they taught and to share their experiences. These meetings supported teachers, such as Adam, for whom the course and pedagogy felt like "uncharted waters." Adam listens to what his colleagues report as they share "lessons that are successful and what lessons they avoid" and takes mental notes about the tasks he think might be effective. Adam tends to choose tasks that have been successful with other teachers and those that he "heard nothing but good things about." Teachers who were "ahead" of Adam provided insights into which tasks "worked" and which "didn't work at all," informing his own decisions about whether to use a particular task.

The "Remodeling Our Classroom" task emphasizes mathematics content that Adam feels is important for students to learn and apply. Specifically, the task involves the use of surface area and unit conversions but addresses these ideas within a real-world, problem-solving scenario. Adam stated that he would use this task not only in MMR, but in other classes as well because the mathematics content knowledge is used by students in a way that is relevant and meaningful for their lives. Such tasks help students to answer the "When am I ever going to use this again?" question that Adam "frequently" hears when teaching other mathematics courses. He believes that, though his students may never remodel a classroom, they might apply what they have learned when remodeling a bedroom or a kitchen. According to Adam, his students “tend to be at ease when they are learning this” compared to when the mathematics is not relevant or applicable to them. The “Remodeling Our Classroom” task and other MMR tasks allow students to “think about and apply it [what they have learned] to their own life.”

Phase 2: Task Planning. The MMR course materials include detailed lesson plans to accompany each Context, including the amount of time it should take to complete (in terms of 45-minute class periods), the goals and objectives for students *and* teachers, and instructional procedures to assist teachers with the recommended pedagogy. Adam finds the lesson plans for the course to be “really helpful” and chose not to deviate from them with the “Remodeling Our Classroom” task. As I described previously, Adam did not modify the task handout or instructions in any way, though spontaneous conversations occurred in class during task setup. Adam also emphasized his use of the supporting materials provided for the MMR pilot teachers:

> If they [the Ohio Department of Education (ODE)] suggest any kind of things that I should focus on before the lesson, any videos that they think I should watch to kind of help me strengthen the lesson or certain things that I need to do on certain days, I need to make sure that I'm doing those.

Adam attempts to follow the prescribed pedagogy for the MMR course and considers what his role as the teacher should be. When planning his instruction, he said, “I go through the instructional procedures and see what is required of me and what is something that I should leave in the hands of the students." Prior to launching a given task, Adam consults the lesson plan for the task and makes sure to be clear about what his "role" is and what the "role" of the students is. However, he noted that "some lessons are more detailed than others," meaning that some of the MMR lesson plans provide detailed, step-by-step instructional procedures whereas others offer more general guidelines (e.g., "implement your launch plan," a phrase that addresses the setup of a task but does not specify what the setup should look like).

The 2020–2021 academic year presented unprecedented challenges for teachers implementing remote and hybrid instruction. Facilitating students' engagement with the "Remodeling Our Classroom" task required Adam to anticipate challenges involving technology and classroom materials that he might not have considered under typical circumstances. Some issues, such as students' internet connections both inside and outside of school, could not be helped. With his task, Adam pondered how to engage his remote students in measuring the dimensions of his classroom without being physically present. "This assignment has to deal with something that is in the room… that they [his students] are not physically in. So how can I help them see the classroom… so that they can actually really be engaged in it?" To overcome this obstacle, Adam had his remote students collect the same measurements as the students in the classroom but had them do so using their own living areas (e.g., living rooms, dining rooms, bedrooms, etc.). For students learning from the classroom, Adam also anticipated what materials they would need to engage with the task. "I have to make sure, do I have the supplies I need?" For the “Remodeling Our Classroom” task, this meant having long enough tape measurers for his students to be able to measure the walls of the classroom efficiently.

Phase 3: Task Setup. Adam sets up MMR tasks by identifying relevant parameters or simplifying assumptions before allowing his students to work in groups. He provides direct instruction “rarely, if at all” in the MMR class, maintaining the instructional practices suggested in the lesson plans despite typically giving “very detailed… instructions going from Point A to Point B in [his] traditional classes.” As described previously, Adam did not begin the “Remodeling Our Classroom” task by telling students how they should do the necessary calculations nor what to include in their proposals. Rather, he communicated the constraints and reminded students that “They have to paint the room. They have to put carpet down. They have to install a gas line and put in a fireplace.” He also helped students to make a simplifying assumption about the walls of the room so that they could determine the amount of paint. Realistically, the lower level of brick lining his classroom required a “special product” to paint, but Adam allowed students to use the same paint for the entire surface area of the walls. From there, students had the freedom to engage with the task in their own way and arrive at their own conclusions. Adam described that “The process with this is to just give them [students] the bare bones… and the instructions and let them drive the lesson.”

Moreover, an important element of Adam’s task setup included the “What do you notice? What do you wonder?” routine that was emphasized throughout the MMR PD and prescribed in the MMR lesson plans. To start the “Remodeling Our Classroom” task, the students were provided an opportunity to “observe the room, notice what they noticed about the room," and ask questions that "sparked because of that they noticed." Adam described this as a general pattern to his lessons: "So what do you notice? Here's the problem and then what questions do you have that spring forth from this noticing?" Adam then followed the next part of the lesson plan and had students estimate the surface area of the classroom as described previously. This stage of the task setup is also important to Adam because it "helped students to get familiar with taking measurements" prior to the data collection piece of the task.

According to Adam, it is important for students to understand how the work they do in class is relevant to their lives, even if some aspects of a task might be unrealistic. "No, we won't have a fireplace in a classroom, but we're inserting that little ripple in order to get them [students] to think about other things that will apply to their lives." Adam also felt that the discussion of the dropped ceiling in his classroom was "necessary" even though it was not called for in the lesson plan because it helped his students to "visualize" where the piping would go. This discussion was also important to Adam because his students probably had not thought about it before; it helped students to learn something new that was non-mathematical in addition to the mathematics that they learned.

Before allowing students to work on the task in their groups, Adam assisted the class by providing a rough sketch of the classroom and a digital illustration of where the pipeline would go using his iPad. Adam assisted his students in this way because "It's hard for them [the students] to conceptualize it" otherwise. This might also have helped students who were learning remotely and unable to physically see the classroom layout.

This part of the task setup differed from the rest, where Adam offered less direct support. However, as stated previously, the students were left to determine the measurements for the pipeline and make the appropriate unit conversions on their own.

Adam dedicated a portion of the task setup to communicating his expectations for students throughout their engagement with the task. This involved "giving them [the students] the rubric and going through it in detail, letting them know exactly, these are the markers, and this is what is expected of them." Adam set the expectation that his students must provide evidence for the cost of materials in their proposals; they could not simply provide a quantity without proper justification (e.g., numerical calculations or a cost found online). The same was true for a previous task that Adam used: "everything had to be detailed. They [the students] had to use a spreadsheet, had to show formulas, had to show… all of their logic and reasoning." According to Adam, setting these expectations for students from the beginning allows students to engage with tasks more readily and contributes to the work they produce as a result.

Phase 4: Task Implementation. Adam facilitated his students' engagement with the "Remodeling Our Classroom" task by observing them as they worked collaboratively in groups. He wanted to encourage productive struggle and allow his students to think and reason through tasks on their own.

> I am starting to see the benefits of letting them [the students] struggle. They [ODE] call this productive struggle, letting them struggle through the problems because you know what it does to the brain, it causes it to grow when they make mistakes.

Adam feels that, by allowing his students to engage in tasks without showing them exactly what to do, they enhance their own problem-solving and reasoning capabilities even if they made mistakes along the way. To prevent himself from limiting students' opportunities to think for themselves, Adam sometimes mutes his microphone and simply observes what students do and say so that he is not the one talking. According to Adam, "as long as they [the students] are making some connections, I don't tend to say anything. I tend to just let them stumble upon and find things."

Part of Adam's role in supporting student engagement involves holding students accountable for participating in groupwork and discussions. To keep students engaged, Adam sets expectations for what students are required to do as they work through each task and reinforces his expectations during his lessons. "No one can just sit back and just not participate," Adam said during the interviews. "Everyone must be talking. Everyone must be either stating what they noticed, posing questions… there's roles that they need to fulfil." Participating and working actively in class are criteria on the rubric for the "Remodeling Our Classroom" assignment, but Adam also made instructional moves to keep students accountable for active engagement. I witnessed this during the first observation day when Adam joined a breakout group that was working on their written proposal. He noticed that one of the students in the group was not talking and asked, "what is your role [in the group]?" When the student answered that he was responsible for working on the gas line, Adam reminded him that the class determined most of that information together and encouraged him to find other ways to help his classmates by asking them what else he could do.

Adam reiterated at the start of each class meeting that his students were required to provide documentation for additional items outside the general requirements and explicate their work in their written proposals. During the interviews, Adam communicated that this was something he does in all his classes and for all his tasks. He wants to see evidence of students' mathematical work and thinking because it is "beneficial" to him as their teacher; he can identify students' mechanical errors himself (e.g., errors in calculations, incorrect algebraic operations, and so on); however, he is more concerned about students' processes: how they proceeded from one idea to the next "logically and conceptually" because these are things that he cannot directly observe from students' work.

Though students could explore and determine what they wanted to include in their remodeled classroom, Adam felt obligated to provide support and suggestions occasionally. For example, when one group was unsure of how to proceed with the paint for the walls, Adam referred them back to the task handout and had them reread it to clarify what they misunderstood. As I discussed previously, Adam also provided a three-dimensional sketch of what he expected for the scale drawing because they still possessed "gaps" in understanding what he was looking for. He realized that, after being asked several questions and noticing that both groups had yet to create a scale drawing, he needed to intervene and provide additional support. When I asked about these interventions, Adam replied that it was something that he did not have a solid outline for in advance. But if he notices that a group is struggling, he expected that others are struggling as well. This causes him to reiterate the expectations for a task and point students back to the rubrics. He says, “go back to the documents, go back to the handout that I gave you. Make sure you’re answering all the questions. Go back to the rubric… see if you are achieving those different things.”

Alternatively, there were occasions where Adam asked questions instead of providing support or clarification. I noticed several teacher questions throughout the observations, including: “How did you figure that out?” and “What questions do you have?” This practice aligns with Adam’s goal of encouraging productive struggle; rather than “jumping in and saving” his students by providing answers, he instead is “jumping in and saying directing questions,” not giving them a complete explanation but pressing them to think further. Adam also asked questions to guide students down a particular path and sometimes repeated what students said in the form of a question to get them to reconsider or evaluate their statements (e.g., “so what are you telling me then, you’re painting over that?”). Adam questions students to help them progress through the task and to clarify what they are saying and doing, even revoicing their own contributions as questions to prompt them to reevaluate their thinking.

Adam typically has students debrief and reflect on the tasks they complete through a whole-class discussion when they are finished. However, I did not observe this part of task implementation due to scheduling conflicts. During this segment of instruction, Adam asks students to reflect on questions such as, “What did you get from this assignment?” and “What are some things that you now know that you didn’t know?” For the “Remodeling Our Classroom” task, he also wanted his students to discuss some of the selections they made that were not provided on the task handout because some are not realistic in a school classroom setting (e.g., a beanbag chair would likely not be permitted because it may be a fire hazard). For Adam, it is also important that students reflect on the real-world aspects of the tasks they complete. Whole class debriefs are also beneficial, according to Adam, because student sharing may cause others to “become familiar with something that they weren’t familiar with before.”

Beth

Beth is an award-winning high school mathematics teacher with 30 years of teaching experience. She has a master’s degree in secondary education in addition to a bachelor’s degree, and the 2020–2021 academic year was not her first year teaching MMR. The 25 students in Beth’s observed section of MMR are all “college-bound” seniors and the majority completed Algebra 2 the previous year, though several others had taken precalculus and had chosen to enroll in MMR rather than continuing to calculus. Because of the COVID-19 pandemic, Beth taught remotely using Google Meet and her two MMR groups met every other day for 50 minutes. My co-observer and I attended one MMR section for 2 consecutive class meetings to observe the same set of students on both days. The IQA and TAG classifications for Beth’s task are presented in Table 6 to provide context for the following two sections.

Table 6

IQA Scores and TAG Classifications for Beth's Task

Score Level	IQA Score	TAG Classification	Mismatch
Phase 1: Selection	3	DM	No
Phase 2: Planning	3	DM	No
Phase 3: Setup	3	DM	No
Setup 1: CF	NS	NC	No
Setup 2: MR	NS	NC	No
Phase 4: Implementation	3	PwC	No
Implementation: Student Work	2	NC	No

Analysis of Beth's Task. The task that Beth selected for observation is titled "Discovering Slope," a student handout from the first MMR Context in a Theme focused on functions. The "Discovering Slope" handout reviews the concept of slope through six stations, encouraging students to use and apply slope in a variety of real-world and abstract mathematical situations. Throughout the following paragraphs, the six stations are referred to as a single task because they address the same mathematical idea (slope) and the use of this idea is made clear to students in the name of the handout. Therefore, students engage in this task understanding that they should apply their knowledge of slope to solve the problems included in all six stations. The stations provide a combination of routine problems that are typical in high school algebra and geometry classes (e.g., finding the slope of a line passing through two given points) and problems involving real-world contexts that might be atypical (e.g., determining the position of a ramp relative to a house that meets the Americans with Disabilities Act [ADA] guidelines).

With a multitude of problems that students must solve, it is expected that the task addresses various levels of cognitive demand; for example, one problem requires students to determine the slope of a line passing through the points (4, 3) and (3, 1) in the *xy*-plane. This process is decomposed for students into three parts: (a) determining the change in *y*, or the "rise," (b) determining the change in change in *x*, or the "run," and (c) calculating the slope, which is assumed to be done by dividing (a) by (b). If considered in isolation, this is a *procedures without connections* problem that would rate as a 2 on the IQA because there is little ambiguity about what needs to be done and how to do it. Moreover, the problem does not require students to make connections to the concepts underlying the procedures being used.

However, another station engages students in an exploration that focuses on generalizing the graphical appearance (or "direction") of the graph of a line based on its slope. This station encourages students to manipulate sliders and investigate the graphical appearance of lines with negative slope, positive slope, a slope of 0, and undefined slope. This station, taken in isolation, would rate as a 3 on the IQA for task potential because the task asks students to identify and describe patterns through an exploration of multiple representations (numerical and graphical). Similarly, other problems or stations would independently rate as either 2 or 3; therefore, the Phase 1: Selection IQA score for this task in its entirety is a 3 based on the highest potential for students to think and reason mathematically throughout their engagement. Though two stations call for explanations, they simply require students to explain the procedures they use or to explain which of a given set of procedures is correct. Therefore, both IQA raters determined that the task does not warrant an IQA score of 4 for Phase 1: Selection.

Beth did not make substantial changes to the "Discovering Slope" task when planning for her instruction. She only made minor adjustments to the handout and the stations to make them suitable for remote learning. For example, Beth created a digital slideshow presentation using Pear Deck, software that allows students to view, write, and draw on individual or group slides. This helped Beth to "see exactly what they [her students] are doing in real time" as they worked in groups to complete the task. Beth also made changes to one station because her students could not see and measure a real-life example of a ramp as they normally would have if they were learning face-to-face. Instead, Beth provided an image of a ramp with its dimensions labeled so that students could still perform the same numerical calculations required by the task. Because the mathematical content and expectations for the task remained the same in this phase, the Phase 2: Planning IQA score for the task is also a 3.

The task was set up with minimal directions given from the teacher and no preliminary whole-class discussion. After spending approximately 15 minutes working on a warmup number-talk that was unrelated mathematically to slope, Beth told her students to begin working and sent them into their Google Meet breakout rooms (groups of 3–4 students each). The cognitive demand of the task remained unchanged throughout this brief setup and, therefore, the Phase 3: Setup IQA score is still a 3. The Setup 1: CF and Setup 2: MR scores associated with the task are NS, indicating that there was no whole-class discussion of the task prior to students starting the task.

My co-observer and I noted students' engagement with the task by joining their breakout rooms and moving from room to room to develop a sense of how the entire class progressed (simulating what we would do if we walked around in a face-to-face classroom setting to the best of our ability). Most students spent a significant amount of class time engaged in problem-solving throughout the various stations but did not frequently communicate their thinking and reasoning about mathematical ideas. Students executed procedures, such as calculating slope using the "rise over run" approach and graphing points and lines. I observed two instances where a student explained how they solved a problem, once to the teacher and once to other students who did not understand what to do, though such instances did not occur frequently.

Overall, the cognitive demand of the task was maintained at an IQA level of 3. The students made connections between the procedure of calculating slope and both numerical and graphical representations. For example, consider a line of slope 3/2 containing the points (5, 9) and (3, a) in the xy-plane, where a is a real number. One group, realizing that a slope of 3/2 indicates a "rise" of 3 and a "run" of 2, noticed that the change in x from 5 to 3 was the opposite of the "run." Therefore, they performed the opposite operation of the "rise" to determine that the value of a was 6 (the result of subtracting the "rise" from the y-coordinate 9). Other groups solved this problem by graphing the point (5, 9) and working "backwards" with the slope of 3/2, realizing that another point on the line was "down" 3 units and "to the left" 2 units because the slope 3/2 indicates a "rise" of 3 and a "run" of 2. This led to the discovery of the point (3, 6) and the conclusion that $a = 6$. Though the solutions used similar reasoning, the first was a numerical approach and the second was graphical. Moreover, the students determined these solution strategies on their own, applying their knowledge of "rise over run" in slightly different ways. Beth did not lower the cognitive demand of the task by telling students what to do and how to do it; instead, she asked questions to illicit and guide students' thinking. For instance, she asked questions such as, "How might you use your strategy to find the slope?" and "Is 12 realistic? Does it make sense?" to prompt students to think deeper rather than simply providing direct guidance and answers.

Though observations of students' engagement with the task suggest that the cognitive demand was maintained during implementation, evidence of students' written work on the task indicates that the cognitive demand may have decreased in some regard. Beth provided six samples of students' work pertaining to the task, taken from what they drew, graphed, and wrote on their Pear Deck slides. Of the 6 samples, 2 indicated high-quality work, 2 indicated average quality, and 2 were selected as "interesting" samples that stood out from the others in the class according to Beth (e.g., one student graphed a single point instead of graphing the three segments representing the side view of a ramp in one station). Most of such samples included only numerical answers with limited evidence of students' thinking and reasoning aside from the calculations that were made to get there. Moreover, students' work on some stations indicates that they still held various misconceptions; for example, one student suggested that the slope formula $m =$

$\frac{y_1-y_2}{x_1-x_2}$ would yield an incorrect result if the order of the coordinates were switched, that is, the formula $m = \frac{y_2-y_1}{x_2-x_1}$ was incorrect. Most students wrote only “rise over run” as justification for written solutions and several instances of student misconceptions were evident. Therefore, the Implementation: Student Work score for this task was rated as a 2.

Beth’s Analysis of the Task. Beth classified the original, MMR version of the task as “in-between” *doing mathematics* and *procedures with connections* using the TAG. Though, when pressed to choose between one of the two, she leaned toward *doing mathematics*. Beth noted that the mathematics in the task was not complex, however, the way in which students were expected to progress through the various stations on their own contributed to the high cognitive demand. Referencing the TAG, Beth argued that the task would involve some level of anxiety for students because they would be unfamiliar with determining the length of an ADA-acceptable ramp, for example. Students could not refer to a predetermined procedure to solve the problems given several of the stations. Beth voiced that some aspects of the task were “more routine,” perhaps typical of what students had seen in the past, such as determining the slope of a line containing two given points. However, she also emphasized that students would not be able to follow procedures mindlessly; the various stations required students to adapt their knowledge of slope and apply it in different ways for each station.

With minimal changes to the task, aside from the adaptations made to fit an online learning environment, Beth also considered the Phase 2: Planning adaptation of the task at the level of *doing mathematics*. She pointed out, however, that removing the measurement aspect of the first station may have made the task “a little bit less…

demanding cognitively because they [her students] didn't have to measure." To maintain this aspect of the task, Beth gave her students a homework assignment prior to our first observation: to measure a ramp where they lived and calculate its slope. Beth felt that students would maintain the level of *doing mathematics* if they did the measuring part of the task at home despite being unable to do it in class. Because the setup of the task was so brief, including only Beth's instructions for students to get in groups and begin working on the task, she also considered this phase as *doing mathematics*.

Beth felt that the task potentially lowered to *procedures with connections* during the first day of implementation. Though students began to achieve some of the standards for *doing mathematics* during the second day, Beth explained that her students were still mostly in the *procedures with connections* level. After the first day of students' work on the task, Beth thought that her students' responses to questions on the Pear Deck slides were "superficial," meaning that their responses were vague and general, lacking specific reasoning to support them. To potentially enhance the quality of work that students submitted, Beth provided feedback at the start of the second day of class and encouraged students to improve their written work from the previous day. She explained that her students' work was "superficial" (i.e., generic responses that lacked depth) and that they needed to "explain their work better."

During the second day, the students met Beth's expectations for working collaboratively in groups, engaging in productive struggle as she frequently questioned them, and developing an understanding of slope as a rate of change. They engaged in "complex and nonalgorithmic thinking, but it took them a little bit of time to get there."

In one instance, Beth provided direct guidance and walked one group step-by-step through the procedure for graphing points in the *xy*-plane. In doing so, she felt that she "might have… lowered it [the cognitive demand of the task] slightly for the group." However, she also thought that she might have raised the cognitive demand for another group who finished most of the stations on the first day because she asked them to go back and find another way to solve the problems and explain the work that they did in writing.

Beth communicated that, overall, her students engaged with the task between the *procedures with connections* and *doing mathematics* levels but inclined slightly more toward *procedures with connections*. She explained, "I heard great conversations… in that upper level," referring to the higher-level categories in the TAG, "but I don't know that they have written that way." Even after providing feedback and reinforcing her expectations for written responses, Beth felt that her students might not have written thorough explanations for how they solved each problem.

Phase 1: Task Selection. Beth's primary focus is student engagement and collaboration. When teaching face-to-face, she preferred to choose activities that allow her students to get up out of their seats and move around the classroom. Her 30 years of teaching has led her to the realization that students do not learn mathematics by "sitting and copying things down and doing an example," rather, she believes that "Students learn by doing." Beth used the "Discovering Slope" task during the 2019–2020 academic year and appreciated how the stations allowed students to move around the room, something that she missed during the 2020–2021 year. However, the pandemic has not prevented

Beth from focusing on student collaboration as one of her goals for students. She feels that the various slope stations are “relevant” and “thought-provoking” for students and allows them to be “more interactive” than they typically are during online instruction. During our interviews, Beth recalled her early teaching years when she taught slope through direct instruction: “I can remember just getting to the board and saying, here's the slope formula, plug the points in and here's your slope.” Her current stance is that she would rather present the content in ways that encourage students to “investigate… and see what slope really is.” By using the stations, Beth wants her students to be able to explore slope for themselves rather than teaching as telling.

Another reason for using the Discovering Slope task was that Beth wants her students to be able to connect slope to a real-world context. She has asked her students what slope was in the past, and they typically recite the equation of a line in slope-intercept form, $y = mx + b$, and “rise over run,” but she wants them to be able to answer questions such as “What does this mean?” Beth wants her students to see inclined objects, such as hills and stairs, in real-life and think about the slope associated with each. Connecting mathematics with the real world, according to Beth, helps students to remember and understand mathematical concepts more effectively. “They just remember it better because it’s something that they can think about outside the classroom… I think one reason that our kids struggle so much… with math is they don’t see anything outside the math classroom.” Beth said that, specifically, connecting the concept of slope with real-world applications helps students to think about rates of change in context. Though the “Discovering Slope” stations include some abstract, purely mathematical problems, the real-world aspects with ramps relate to the other MMR contexts in the Theme focusing on functions.

Beth's goal was for students to explore the concept of slope throughout their engagement with the task. The same week that I interviewed her, she met with the ODE writing team and they had discussed removing the slope stations from the MMR course. Beth was "pretty adamant" that the task was important for students because many of them still struggle with slope in precalculus and even in calculus. Beth and I heard students in her MMR class state that they were "really bad" with slope, further emphasizing the importance of the task and the content it addressed. Moreover, Beth feels that the task "lays the groundwork for linear functions," which are addressed through numerous tasks in the MMR course. By spending more time with slope, Beth thinks that her students will be more prepared to engage with tasks involving linear functions because of the mathematical connections between the two.

Beth decided to change the order of two MMR tasks: the MMR lesson plan for the "Ramps" Context includes three tasks, given in this order: (a) "Ramp It Up" (taught by Isabel and explained in another section), (b) "Discovering Slope", and (c) an activity in which students design their own ramp. However, Beth chose to rearrange the order of the first two tasks so that her students could do the "Discovering Slope" stations first. This decision was made simply because Beth found it easier to transition "Discovering Slope" into an online activity using Pear Deck and it fit better for her plans that week. Like Aaron, Beth also communicated that "There's only a few lessons from each theme that I can do strictly online easily." The "Discovering Slope" task fit the criteria and therefore was selected as a task that Beth would use.

Phase 2: Task Planning. Most of Beth's planning involved transitioning the stations into something that students could do online rather than face-to-face. "Normally, when we are in person, I don't have to do much at all to alter it [the task]… but online, I did have to turn it into something they could do online." For this purpose, Beth used Pear Deck because the software fit her needs as well as those of her students. Pear Deck is a software that allows students to view and edit presentation slides, created by a teacher, collaboratively. Teachers can view students' work live and provide feedback as they progress from one problem or task to another. Beth appreciates the ability to observe and monitor students' work using the technology: "With Pear Deck, I can see exactly what they [her students] are doing in real time." Because Beth's students were not required to turn on their webcams when attending class online, she felt assured knowing that they were participating and engaging by what she could see using the software. "Pear Deck is the only way that I can tell what's going on… that's the only way I know they are with me or not… I could not do online instruction without it." Pear Deck also allowed Beth to provide her students with drawing tools and digital graph paper on additional slides that they could use to help them work through the stations. "There's a couple of the activities that ask them [the students] to draw graphs, and it [Pear Deck] is the only way that I can get the kids to draw a graph that I can see electronically." In general, planning remote instruction requires Beth to be "much more deliberate" than she had been before. Each task forces her to consider the technology that might be most appropriate based on what the task requires of her students.

Planning around her students' interaction and engagement with the task required Beth to anticipate (a) how her students would work collaboratively and (b) how they would respond to various parts of the task. In terms of groupwork, Beth indicated that "some groups… are better than others online." Because one of her goals for the course is for students to participate collaboratively, Beth monitors their interactions with each other and noticed that some students do not communicate well with each other. Anticipating students' interactions and organizing groups accordingly is an aspect of planning that might not appear to influence the use of a task. However, taking preemptive action to promote students' teamwork and collaboration can support students' engagement with a task and the quality of mathematical thinking and work they achieve. Beth's strategy of anticipating student responses informed her planning because it helped her to determine appropriate questioning strategies in advance. To help students visualize a slope of 3/2 as a change in y over a change in x, as opposed to thinking of it as 1.5, Beth predicted that she would ask students questions such as, "What does a slope of 3/2 mean?" Asking questions such as this, she thought, might help guide students toward thinking about a slope of 3/2 as a "rise over run."

Phase 3: Task Setup. The setup of this task was minimal from Beth as she simply indicated for her students to enter their breakout rooms and begin working on the task. However, Beth posted her expectations on her Pear Deck slides for students to read as they did so. Her expectations included that all students were to actively participate vocally using their microphones and that each group worked together on the same problem at the same time so that no individual students moved ahead of their group members. These expectations help Beth to set the tone for how her students should engage in a task, but she also reflected that, “There were some things I could have done to be more specific about the directions” to ensure that students were able to locate their materials and get started on the task more efficiently on the first day. During my first observation with Beth and her students, I also noticed that many students needed a reminder of what they should be doing and where to find the online materials. “Wednesday, I wasn’t really happy with how it [the task] started,” Beth explained. “I think I wasted fifteen minutes getting everybody where they were supposed to be with the right material, on the right work, and everything.” She indicated that, in the future, she would include this information on her Pear Deck slides to help students get started with their work for the day and make the setup phase of future tasks more efficient.

Phase 4: Task Implementation. This is not the first time that Beth has taught the MMR pilot course and implemented the “Discovering Slope” task. With each implementation, she has reportedly allowed her students to engage with the task more and provided less direct support. “The first year, I tended to… help them more,” Beth explained. She took the same approach this year, supporting students’ mathematical thinking and reasoning by questioning and providing suggestions to guide students without taking away from the work of the task. In one instance where a group had miscalculated the slope of a ramp, Beth asked questions to help them see their mistake on their own. The students had erroneously divided 29/348 and claimed that the result was

12, to which Beth wrote $\frac{29}{348} = \frac{12}{1}$ on a whiteboard and asked, “Is this what you mean?” This led one of the students to state that 29/348 should result in a number less than one and then realize her mistake. Beth asked questions such as this because she “wanted to get students to 1/12 on their own” rather than telling them what the result should be.

Beth felt that some students were “too stuck on the formula” for calculating slope, leading them to have trouble with the first station because it gave the dimensions of a ramp rather than coordinate points. The cause for some students’ difficulty, according to Beth, was that they could not identify the numbers y_1, y_2, x_1, and x_2 to “plug in” to the formula $m = \frac{y_1 - y_2}{x_1 - x_2}$ because only two numerical dimensions were given. To help students determine the slope of the ramp, Beth asked, “What is another way of saying slope?” The students were able to recite the formula back to her, but that made Beth realize “that they weren’t clear on what the meaning of slope was… that’s why I kept trying to get them to go back to the rise over run, to try to get them to start there,” she explained. Beth asked her students to think about other ways of describing slope because she wanted them to understand it as more than a formula and a procedure. She also suggested that multiple groups try to “draw a picture” to help them visualize what was being represented in the task and think about slope graphically.

When Beth identified a common student misconception, she began to check students’ work frequently to diagnose whether this mistake occurred in other groups so that they were not making the same mistakes. She noticed during the second day of class that some of her students consistently plotted (x, y) with the coordinates reversed, graphing the x-value as the vertical coordinate and the y-value as the horizontal. Beth described, “Once I saw one group had graphed the point backwards, I started… actively checking for that,” correcting students if they had it wrong. Afterward, Beth wondered if her students understood that the first coordinate in any (x, y) pair corresponds to a value along the x-axis and the second coordinate corresponds to a value along the y-axis; it seemed to her as if her students only “memorized the process” of moving in a direction from the origin (i.e., up, down, left, or right) and did not associate the numerical coordinates with the axes.

As stated previously, one of Beth’s goals for students was for them to be able to explain their thinking and reasoning in words. Correct answers were not enough. To support students in developing this skill, Beth comments on their work frequently and allows them to resubmit assignments after incorporating her feedback. This practice was evident at the beginning of the second observation when Beth asked her students to improve some of their “superficial” responses. For example, she wanted students to provide more elaborate reasoning for their answers than, “because we graphed it.” Instead, she sought a reason for how her students found their answer using a graph.

Beth encouraged students who had completed the stations early to go back and try to solve the problems using different approaches. Even for a specific student who Beth knew could do the calculations mentally, she asked him to “write out” how he found the answers. Beth’s stresses providing written justification is because it forces students to “think about how they can explain it [their ideas] using the right vocabulary that someone else could understand.” It also “makes it more in depth than what they [students] are used to doing.” However, Beth explained that it has become more difficult to get students to communicate their thinking when teaching remotely: “In person, when they try to explain, I think they have rich discussions. It’s not happening as much online as I would like.” One of the challenges for Beth during the 2020–2021 academic year has been to foster student discussions and explanations in the same manner as she had before when her students met face-to-face.

Cathy

Cathy is in her 26th year of teaching, all at the same high school. She has earned a Bachelor of Arts in mathematics, a Bachelor of Science in secondary education, and a Master of Education in secondary education with coursework toward a PhD as well. The 2020–2021 academic year was Cathy’s first year teaching the MMR course and she also attended the ATC Modspar professional development program hosted at Ohio University for one summer. There are 28 students enrolled in her MMR course, 10 in one section and 18 in the other (the observed section); 2 are juniors and the rest are seniors. My co-investigator and I observed her group of 18 students for 3 days because Cathy suggested that it would allow us to observe the entire implementation of a task and there were no scheduling conflicts preventing us from doing so. Cathy teaches face-to-face, though some students learn remotely because they were quarantined to prevent the spread of the COVID-19. Remote observations with Cathy were conducted via webcam the same way in which remote students accessed the course from home. The IQA and TAG classifications for Cathy’s task are presented in Table 7 to provide a summary of the following sections.

Table 7

IQA Scores and TAG Classifications for Cathy's Task

Score Level	IQA Score	TAG Classification	Mismatch
Phase 1: Selection	3	DM	No
Phase 2: Planning	3	DM	No
Phase 3: Setup	3	DM	No
Setup 1: CF	3	NC	No
Setup 2: MR	2	NC	No
Phase 4: Implementation	2	DM	Yes
Implementation: Student Work	3	NC	No

Analysis of Cathy's Task. The task that Cathy used for this study, titled "Starbursts™ Grab," was provided for MMR teachers with the course materials. The task is meant to engage students in learning about linear regression as they design an experiment to predict how many Starburst™ candies they can take from out of a bowl with a single grab. According to the MMR lesson plan, students are expected to determine and measure the variables of interest, collect data by performing numerous grabs, and use linear regression to develop a prediction model for the situation. Students are meant to represent data in a spreadsheet and explore the possible correlation between the size of a person's hand and the number of Starbursts™ that they can grab.

The "Starbursts™ Grab" task scored as 3 for Phase 1: Selection because the task involves nonalgorithmic thinking and students are not provided a predictable, well-rehearsed approach to complete it. For example, students must determine how they will measure and collect data for hand size on their own. Moreover, students are not given a method to determine whether their linear models are accurate predictors for the number of candies a person can grab. Students are also expected to identify patterns and form generalizations using the data they collect and analyze. Students work with and interpret data in various forms throughout the task, including numerical data, graphs, and algebraic equations. However, the task does not explicitly prompt for evidence of students' reasoning and understanding; the questions on the handout do not encourage students to explain how they interpret their data and how they make their conclusions. For example, the question, "Does there seem to be a correlation based on the evidence you see here?" asks only if students can identify the presence of a correlation, not to explain how they know a correlation is present or the reasoning behind their decision.

The task scored a 3 on the IQA for Phase 2: Planning for the same reasons as described earlier. Cathy decided to have students complete the task individually to avoid unsafe contact among students, but otherwise the wording on the task handout was generally unchanged. Some statements had the collaborative elements removed, and other changes involved minor rewording and splitting larger sections of text into multiple, smaller ones. The handout was also enhanced in some ways; for example, the statement, "As a class, decide which variable you want to use. List those below," was changed to "Describe the variables we chose as a class and how we will measure them," suggesting that Cathy's version of the task required more depth in terms of students' written work in some respects. However, the questions that students were required to answer still did not explicitly prompt students to explain their reasoning or justify their answers. Cathy created three slides to help facilitate her students' engagement throughout the task, one for each day, but the slides only included broad instructional procedures and did not lower the cognitive demand of the task. Such instructions included prompts for students to post their "notices and wonderings" online, a request for a volunteer do make the first candy grab for the whole class to observe, and directions for students to answer a specific question on their handouts.

Phase 3: Setup, task setup, scored a 3 on the IQA rubric for task potential as well; a preliminary whole-class discussion occurred prior to students' work on the task, but the cognitive demand of the task did not decline. Task setup began with students making two predictions, (a) the dimensions for a Starburst™ candy and (b) how many Starburst™ candies a student could grab out of the large bowl positioned in the front of the classroom. Students provided several estimations for the dimensions of a single candy: one said that the square face was 1-in by 1-in and each candy was half an inch thick; another thought that the square face had an area of 1.5 cm^2 and was 1 cm thick. The students also offered several predictions for how many candies they could grab out of a bowl, ranging from as few as 11 to as many as 35.

After discussing their answers as a class, Cathy had her students post what they noticed and what they wondered about the bowl of candy on an online social-networking site called MeWe. Students' observations included that the bowl was almost half full, that there were four colors of wrappers, and that there were "lots" of candies in the bowl. Examples of what they wondered include how many candies were in the bowl, how many candies it would take to fill the bowl, and whether the number of candies in the bowl affects how many can be grabbed with a single grab. Cathy elicited what students knew about the real-world context of the task: they discussed the technique used to grab the Starburst™ candies (using either a "scoop" or "claw" method), how hand size varies among people, and how the activity reminded them of games that they had seen where the objective was to guess the correct number of candies in a jar. Multiple students participated throughout these discussions, but ideas were typically expressed in isolation and not connected to each other; therefore, the Setup 1: CF score for task setup was designated as a 3.

As a class, Cathy and the students also discussed potential variables of interest and how they planned to measure them. Some examples that students suggested were the tightness of the person's grip, their hand size, and the number of candies in the bowl. They also determined whether the potential variables could be measured, as well as the parameters that could change and those that would remain constant. Cathy asked what her students thought about each potential variable as they progressed through the list they had generated as a class; for example, she asked what students meant by the "grab technique" (scoop or claw method) and whether it was something that would remain constant throughout the experiment. The use of accountable talk moves was infrequent at best, but students actively participated in the discussion, warranting a Setup 2: MR score of 2.

The setup of the task occurred during the first observation, whereas the implementation of the task took place during the second and third days. Students spent the second day individually collecting data as Cathy monitored their progress from a safe distance. The students entered their data into individual spreadsheets and Cathy combined them all into a class spreadsheet for further analysis. The teacher made several suggestions throughout the process, indicating that students should measure the size of their hands to the nearest quarter inch and that they should enter only numerical values into their spreadsheet and type the units in the column headings. The third day involved having students use technology to calculate a linear regression equation for the class data and using it to predict how many candies Cathy would grab based on her hand measurements. Both Cathy and a guest teacher took measurements and performed three grabs from the large bowl.

The analysis that followed, however, was done mostly by the teacher with limited student interaction. For example, Cathy told students to "Plug my measurement into the formula, see what you get," after which she concluded "I don't think it [our model] is the best model." Her directions to students consisted only of procedures to execute and students did not verbally communicate their understanding about linear regression and making predictions. It was Cathy, rather than her students, who stated that the thumb-to-pinky measurement appeared to be a weak predictor for the number of candies that were grabbed. Because students simply performed calculations using their regression equations and did not connect the results back to the real-world context, the data, or other representations, the Phase 4: Implementation IQA score for this task is 2.

Cathy submitted 14 samples of students' work on the "Starbursts™ Grab" task, and I selected 6 to be analyzed using the IQA: 2 were selected as examples of high-quality work because their writing demonstrated an understanding of how their linear model related the size of a person's hand to the number of candies they could grab. One student wrote that "If your hand size is between 7–9 inches then you're likely going to grab between 20–30 starbursts," suggesting that they understood how to interpret the model and its associated graph to make predictions. The 2 samples of average student work contained numerical calculations but only limited written explanations; typical of many students in the class, the students answered some questions on the handout with only "yes" or "no" responses. The remaining 2 samples were unique because the students contradicted themselves in their own responses; for example, one student noticed that "bigger hands grab more candy" but also wrote that none of the variables measuring the size of a hand could be used to predict the number of Starburst™ candies that were grabbed.

Overall, the six samples demonstrated that students engaged in some level of problem-solving as they described the variables they explored, generated linear regression models, and interpreted the meaning of their results. Across the student work samples, students identified a pattern that the number of candies increased with hand size; however, they failed to provide strong written explanations for how they came to this conclusion. The students used numerical, graphical, and algebraic representations but did not communicate explicit connections between them in their writing. Therefore, the Implementation: Student Work score for this set of samples is 3.

Cathy's Analysis of the Task. Cathy classified the original MMR-version of the Starbursts™ Grab task at the level of *doing mathematics* in the TAG. She explained that the task requires students to engage in complex, nonalgorithmic thinking as they determine their own variables and ways of measuring them. The use of spreadsheets, according to Cathy, involves some algorithmic thinking but does not lower the task to *procedures with connections*. She stated that the task requires students to explore and to understand the nature of mathematical concepts, demands self-monitoring of one's own cognitive processes, and requires students to access relevant knowledge and make use of it; however, she did not provide further explanation as to why she thought so during our 1-hr timeframe. However, she explained that students must analyze the task and actively examine task constraints because they could choose how they want to measure hand size and determine which measurement yields the most accurate predictions.

The task at the planned phase was still at the level of *doing mathematics* in Cathy's mind. She stated, "I don't think I took away any of the thinking… any of their need to create or apply… in the changes I made." She acknowledged changing the activity from group to individual work, but that the class would decide the variables to measure together. Cathy felt that these changes did not lower the cognitive demand of the task because they did not change the mathematical work that her students would do.

Cathy expressed that the cognitive demand of the task may have lowered slightly in the setup phase but stayed at the level of *doing mathematics*. Though students engaged in complex, nonalgorithmic thinking by selecting and measuring their own variables, Cathy thought that she might have directed her third period class more than she wanted to because she wanted to be able to combine her second and third period data together. In terms of analyzing the task and examining task constraints, she reflected, “I might have taken a little of that away in that discussion where we decided with which variables we would control, and which ones were constant… I think I guided that discussion more than I might have.” Cathy felt that she needed to guide this segment of instruction because she did not have the time to allow students to make these conclusions on their own.

Task implementation fell between *doing mathematics* and *procedures with connections*, Cathy considered. Students were still using nonalgorithmic thinking by “thinking out of the box,” using their own thinking rather than limiting themselves to using a particular formula or strategy. Cathy also indicated that she did not instruct students on how to determine whether a regression equation was a “good predictor,” but students were able to make sense of it on their own. Though students were expected to write about this on their handouts, it was Cathy who voiced that the model was not a “good predictor” in class. She summarized her view of the lessons in the following statement:

> I think they got an idea of what it means to look at a situation and analyze it and do the math, I really do… This isn't the old fashioned, ‘I'm going to give you every step of the way to go.’

However, she also noticed some aspects of students’ engagement suggesting that the cognitive demand of the task declined. For example, she felt that she and her students could have done more to explore and understand the nature of mathematical relationships because they did not debrief via a whole-class discussion after completing the task. The

MMR materials "didn't give [her] a good enough post discussion" but she considered spending additional time to debrief as a class on another day. She also stated that "based on their responses, I don't think they [the students] stretched themselves as much as I would have liked… they might still be more at a *procedures with connections* level as far as their thinking." Cathy concluded that, despite there not being "much struggle left" in her students' engagement with the task, she "kept more of the bullets" in the TAG than she "lost," leading to the result that her students were still at the *doing mathematics* level.

Phase 1: Task Selection. Student engagement is Cathy's primary focus when selecting tasks to use with her students. Engagement has been her focus from her first year teaching, as she was influenced by teacher mentors and others involved in OCTM. She learned the phrase "never say anything a kid can say" by reading an article published in NCTM's *Mathematics Teacher* journal and tries to select tasks that engage students in developing meaning through exploration. The MMR course has provided Cathy with access to a wealth of materials centered on active student engagement and she has enjoyed every day teaching it. Cathy explained,

> This has been an entire year where every single lesson is my favorite, lessons that occur a couple of times a quarter, maybe, in other courses. So once or twice a quarter, if I'm lucky, I manage to squeeze in something that I really love. It's really cool that students are really engaged and that I think it's amazing.

The MMR course has allowed Cathy to use engaging, explorative tasks more frequently than with other courses because the materials were provided to her; she has not needed to find or adapt the materials herself. She explained that she does this for other courses, but “the problem is the time it takes to take something and convert it.” It takes her sometimes several hours to adapt a task from source materials into a form that she wants, and then it may take several years of refinement after each successive implementation for the task to reach its final form. Cathy is “thankful” for the MMR course because the materials are provided and much of this work is already done for her.

The MMR course has also enabled Cathy to use tasks she would not have considered before due to the instructional materials required. She would have “never” chosen the “Starburst™ Grab” task prior to teaching the MMR course because she would have been required to spend her own money on materials. Material-intensive tasks make Cathy hesitate because she feels she cannot afford to purchase new materials for each one. However, she has been able to use material-intensive tasks because her district offers financial support while she pilots the MMR course. The district purchased the candy she needed for the observed lessons along with a class set of Barbie dolls for “Barbie Bungie,” an MMR task that engages students in developing the optimal bungee cord made of rubber bands. Though both were tasks that Cathy was interested in, they would likely have not been used if she had not been provided the resources to make them possible. The MMR course has encouraged Cathy to be “willing to go out on a limb and try these things that are material-intensive” that she might not have done otherwise.

The mathematical content and processes addressed by the “Starbursts™ Grab” task are also desirable features that Cathy highlighted. She appreciates that many MMR tasks allow students to gather, represent, and interpret data using spreadsheets and regression models. According to Cathy, such tasks help students to see how these processes are both useful and applicable in the real world and in purely mathematical contexts. Engaging in data collection, management, and analysis in a task is helpful, but through numerous tasks helps students to realize that these are skills that they can really use in other contexts. Specifically, the use of spreadsheets is something that Cathy has encouraged for many years. “We really need to teach kids how to use spreadsheets,” she said. “If any department in the building is responsible for spreadsheets, that would be the math department. I’ve said that for years. We’ve never done anything about it, but I’ve always thought that.” Through “Starbursts™ Grab” and many other tasks in the MMR course, Cathy believes that students are learning spreadsheet skills that will be useful to them in the future.

Phase 2: Task Planning. The MMR tasks are meant to be completed by students in groups. However, Cathy modified the task so that students would complete it individually to avoid unnecessary health risks due to COVID-19. Having students work individually meant that they would not need to share materials and resources and would ensure that students remained socially distant. Aside from potentially requiring other students’ help to measure their hand size, Cathy thought that the task could be completed individually and students would still do the same mathematics. Another added benefit of an individual task was that students who were “fatigued” from group work could get a break from it. Cathy stated that “kids who tend to pick up the slack all the time” might benefit in this way. Designing the task for individual work required Cathy to remove a question from the MMR handout. The question involved analyzing students’ group data, which would not make sense if students collected data on their own.

Cathy also modified the task to accommodate students learning remotely. Remote students were not able to participate in collecting data by grabbing candy from the bowls in the classroom. Instead, Cathy assigned them some problems to solve and allowed them to use the combined class data once they were gathered. The three problems for remote students involved finding the dimensions of a Starburst™ candy and rearranging various numbers of them into rectangular prisms to calculate the volume of. The modified problems are not directly related to the linear regression task, but Cathy felt that it was difficult to make remote data collection “meaningful.” She contemplated the notion of asking students to grab candy from bowls they had at home, but this raises various issues. Cathy identified that the sizes of candy and bowls may be vastly different from those in class and would not contribute to the class data set. Even if they could, the students would not have enough data to calculate a reasonable linear regression. Though this option was suggested by other MMR teachers in one of their meetings, Cathy chose to assign other problems and allow her remote learners to use the full-class data instead. “Everybody now has the class data, and everybody does the mathematics,” she said.

Generally, Cathy tries to plan around the original intent of the MMR course developers when using the course materials. Aside from making the discussed changes to make a task suitable during the year of the pandemic, she looks to what the MMR materials suggest and tries to “stay true to them, whatever they [ODE] want the lesson to look like.” The MMR teachers would normally be somewhat limited in what they could do when modifying the instructional materials for the course. However, the 2020–2021 academic year was different because the COVID-19 pandemic radically altered the way that many teachers teach and the way that many students learn. No data were collected in the fall of 2020 for the MMR course, but Cathy tried to be "as true to the materials" as she could be regardless.

Cathy modified the presentation slides she developed to accompany the task handout between each lesson. These changes were made as she reflected on students' work on the task each day, including reminders to the students and reminders for herself. For example, one reminder Cathy added on her slides for Day 2 was a reminder for the class to talk about precision. During our interviews, she explained that her second period MMR class had decided to measure hand size to the nearest quarter inch; however, this discussion did not occur naturally during her third period class (the class I observed). Cathy included the reminder on her slides so that her third period class would have that discussion and determine the level of precision they would use when collecting data. Cathy also included the constraints that the class agreed on when collecting data (e.g., "Grab with the intent to get as many as you can") because she felt that students would forget them the following day. This form of planning stemmed from Cathy's reflection on the previous day's class and anticipation of what she would need to do to prepare her students for the following day.

After implementing the task, Cathy reflected that one thing she should have changed was to reword the "yes or no questions," questions on the handout that did not ask students to provide explanations and evidence of their reasoning. Astounded, she asked, "Can you believe that this handout had 'yes' or 'no' questions? I even had that conversation with you about how I never give them [students] 'yes' or 'no' questions."

She noticed that, after reading through students' written responses, many students simply wrote "yes" or "no" in response to questions such as, "Does there seem to be a correlation based on the evidence you see here?" Cathy noticed that the wording of the question did not prompt students to provide reasoning for their answer and acknowledged that she did not blame the students; they answered the questions as they were written. After realizing this, Cathy made changes to the document so that the questions would require stronger written responses from students for the following year.

Phase 3: Task Setup. The setup of the task began with the "What do you notice? What do you wonder?" routine that is typical of MMR tasks. For this specific task, Cathy had her students post online what they noticed and wondered about the large bowl of Starbursts™ that she positioned at the front of the classroom. The decision to do this was, in part, motivated by the instructional procedures emphasized in the MMR materials and Cathy's desire to "stay true" to them. However, she has also practiced a similar technique prior to teaching the MMR course in the form of "stand and talks," where she has students rise out of their seats and share ideas with other classmates in pairs. "It gets them [her students] thinking, they're engaged," Cathy stated during our interviews. Due to the health risks associated with COVID-19 during the 2020–2021 school year, Cathy facilitated this portion of instruction using MeWe, as I described previously, because it provided a safe way for students to communicate their thoughts and questions about the Starbursts™ online.

Prior to collecting data, Cathy also led her class in a discussion of the potential variables of interest her students had identified through their noticing and wondering and how they might measure them. The class also discussed which parameters would remain constant throughout the experiment (e.g., using the "claw method" for every grab). Cathy felt that these were discussions that "we don't have often enough in math class" because typically these decisions are made for students in advance. Cathy explained that traditional "word problems" provide students with the assumptions and the data, prohibiting such conversations from happening as often as they should. As I mentioned earlier, however, Cathy appeared to dominate this segment of instruction rather than allowing her students to do the decision-making. Interestingly, she made the same reflection during her post-observation interview: "Since we hadn't talked about it much, I did guide that one a little more than I would if I had… if I wasn't worried about whether we'd have enough time." Cathy later repeated that "I think I guided that discussion more than I might have… I was worried I wouldn't have time."

Cathy frequently sets up tasks by having students make predictions or estimations; this is true for MMR tasks but also tasks that she creates or modifies herself. For the "Starbursts™ Grab" task, Cathy did this by posing two problems for students to consider as they entered the classroom on the first observation day: the first was to "Estimate the dimensions of a Starburst™ candy" and the second was "How many Starburst™ candies do you think you could grab out of a large bowl full?" Cathy asked these questions to keep her students busy so she could take attendance as they entered the classroom, but also to get them thinking about the task they would complete over the next 3 days.

Phase 4: Task Implementation. Cathy felt that the wording of the questions on the Starburst™ task handout influenced students' implementation of the task. She wants her students to be able to explain their work and reasoning, highlighting that "describing their work" and "putting it into words… is a great skill to have." However, Cathy reflected that many students simply provided 1-word answers of "yes" or "no" on the handout because that was all that was asked of them. She shared several examples during our post-observation interview, reading students' responses aloud: "Could any of these variables be used to make predictions? Yes… They answered the question. I can't argue with you, you answered the question." Cathy was surprised that she allowed those questions to remain on the handout but did not blame students for how they responded. She planned to bring this up with students the following week, saying "That's on me. I asked a 'yes, no' question." However, she also planned to ask students, "Let's think about it a little more… What would you have said if I asked this? Let's think about a better way to answer that question." Cathy believed that her students did not provide rigorous responses in some instances because the questions themselves were not rigorous; however, she also felt that she could enhance the outcome of the task by improving her questioning and following up with students.

As I described previously, Cathy did not conclude the implementation of the Starbursts™ Grab task with a full-class discussion or debrief. During her post-observation interview, Cathy explained that the rationale for this decision was primarily based on her experience with remote and hybrid learning. She confessed that she struggled "standing in front of the room and asking questions" because "you get one or two people that might talk, and with half of them [her students] at home, and now in some classes more than half of them at home, it's just not engaging." In fact, more than half of Cathy's MMR students attended class remotely during the 3 days of the "Starburst^{TM} Grab" task (though many were in the class that I did not observe). Though Cathy wore a headset and microphone so that she could communicate with her remote learners and reminded them that they could unmute themselves and speak, she has experienced difficulty getting them to talk during class. Cathy thinks that this is in part because they are "not comfortable" speaking up when learning remotely. She described that she has the same difficulty when leading and attending remote conferences and professional development. These experiences have led her to the conclusion that "Full-class discussion just doesn't work" in the online learning environment because "if grown adults are intimidated to unmute and talk, how in the world can we expect teenagers to feel comfortable doing that?"

Remote and hybrid instruction provide additional challenges that Cathy considers when implementing tasks. Holding students accountable for actively engaging in class is one such challenge that she struggles with. Cathy sometimes provides comments when students are asked to type written responses online by saying "Hey, why aren't you typing anything there? What do you think?" However, monitoring remote students' engagement while she is occupied with her face-to-face students is more difficult. Some students, Cathy explained, attend class remotely but position their webcams so that their faces cannot be seen (I noticed this during my observations). Others, according to Cathy, make themselves visibly present but "mute me and aren't paying attention to class." She tries to hold students responsible by threatening to mark them absent if she notices this behavior, but she does not have the time to constantly verify whether her remote students are engaged in class activities. Task implementation, Cathy reflected, can be a struggle when her students are not in the classroom with her.

Debbie

Debbie's education includes a bachelor's degree in mathematics and in chemistry, a master's degree in physical chemistry, a Master of Education with her teaching license, and a master's degree in educational leadership. She has taught mathematics and science for more than 20 years, and the past 3 years have been at her current school. The 2020–2021 academic year was Debbie's second year teaching MMR. There are 16 students taking the MMR course at Debbie's school, 7 in one section and 9 in another. Debbie teaches the whole group of 16 remotely on Mondays for 30 min, one section face-to-face on Tuesday and Thursday for 41 min, and the other group face-to-face on Wednesday and Friday for 41 min.

My co-observer and I joined the Wednesday-Friday class of 9 students remotely via Zoom; a laptop equipped with a webcam was positioned on one side of the classroom so that we could see and hear as much of each lesson as possible. We chose to observe this group because there were more students, and they were "more interactive" than the Tuesday-Thursday group according to Debbie. One student was absent on both Wednesday and Friday, a different student each day, resulting in a total of 8 students observed during each class meeting. The IQA and TAG classifications for Debbie's task are presented in Table 8 to provide context for the next two sections.

Table 8

IQA Scores and TAG Classifications for Debbie's Task

Score Level	IQA Score	TAG Classification	Mismatch
Phase 1: Selection	4	PwC	No
Phase 2: Planning	4	PwC	No
Phase 3: Setup	4	DM	No
Setup 1: CF	3	NC	No
Setup 2: MR	3	NC	No
Phase 4: Implementation	4	PwC	No
Implementation: Student Work	4	NC	No

Analysis of Debbie's Task. Debbie, Ethan, and Fred selected the same task to use for the purpose of this study, the first part of a 2-part MMR Context titled "Follow the Bouncing Ball." The associated student handout for the first part of the Context is meant to guide students through an exploration of the relationship between the height at which various balls are dropped and their returning bounce height. Throughout the task, students explore this relationship by identifying independent and dependent variables, predicting what type of functional relationship they might have, performing an experiment to collect data, and interpreting the results. The data are expected to demonstrate the linear relationship between rebound height and drop height, and therefore, students also interpret correlation coefficients and use their model to predict rebound heights from various drop heights. An interesting feature of the task is that, though students are initially asked to identify their own variables for the first two on the handout, the next question directs them to use drop height and rebound height as their two variables throughout the remainder of the task. This seems counterproductive and limits the ability for students to explore, though it also helps teachers to focus on aspects that will yield relatively linear data.

I scored the student handout for this task as a 4 for Phase 1: Selection using the IQA. The task explicitly prompts students to make mathematical connections by identifying independent and dependent variables that might have a functional relationship. Throughout their engagement with the task, students also make conjectures by sketching a graph to illustrate such a functional relationship. Students are expected to make mathematical connections between various representations, including numerical data, models (algebraic functions), and their associated graphs. Students are asked to make these connections through language such as the following: "Based on your scatter plot what functional model do you think will best represent the relationship between the two variables you are experimenting with?" Students identify patterns among their data through questions such as, "Does there appear to be a pattern to the data?" This task warrants an IQA score of 4, rather than 3, because students are prompted to explain their reasoning and to show their work numerous times on the handout.

Debbie generally left the task unchanged in her planning and made only minor formatting changes to the task handout, such as splitting large paragraphs into smaller portions and adjusting the syntax of the written text. She also broke larger questions into smaller ones so that the total number of questions was greater in her version, but the expected outcomes from students remained the same. The only influential change to the task handout was the inclusion of a step requiring students to organize their thoughts prior to collecting data. The added section prompts students to provide a hypothesis for what will happen to the rebound height throughout the experiment and has students complete a data table. This section also prompts students to identify aspects of the experiment that would remain constant "to ensure consistency." However, the mathematical nature of the task remains the same with these changes to the task handout, as the language for each question is nearly identical to the original MMR version. Therefore, the Phase 2: Planning IQA score for the task is also 4.

Phase 3: Setup, task setup, remained at IQA level 4 because Debbie's students did the mathematical thinking, reasoning, and communicating during the initial discussions prior to task implementation. The whole-class discussion that occurred included both the contextual features and the mathematical relationships involved in the task: the first part focused on the 1965 Super Ball television commercial that Debbie played for her students in class. She asked the students if they had something like the Super Ball and if any of the students had seen one before. Some of them had seen something similar, though not exactly the same thing, and others were curious and researched them using their laptops. Debbie and her students made connections between ideas that helped to support students' understanding of the real-world context (albeit inconsistently), resulting in Setup 1: CF score of 3.

The discussion of mathematical relationships focused on identifying measurable features of a ball that might influence its bounce and determining which could be appropriate dependent and independent variables. For example, students suggested that the drop height, the size of the ball, and the material that the ball was made from may influence the ball's bounce. The students provided their own ideas throughout this portion of instruction and the use of accountable talk moves was consistent, yielding a Setup 2: MR score of 3. One example of Teacher Press, for instance, occurred after Debbie asked her class why it was important for some aspects of the experiment to remain constant throughout each trial. One student responded that this would yield accurate measurements, to which Debbie followed up with "Why?" This press forced the student to consider how consistency in measurement and data collection affect the results of an experiment.

The observed implementation of the task, Phase 4: Implementation, also received an IQA score of 4. After students worked in groups to collect data, Debbie facilitated their engagement by (a) asking them to first talk to each other about how they would answer each question on the handout, then (b) instructing students to write their individual responses, and (c) bringing the class back together to discuss what they had written; this process occurred in cycles. Throughout task implementation, Debbie explicitly prompted students to provide evidence of their thinking and reasoning. For example, she asked the class to explain if there was a pattern in the scatterplot of the class data and why it might have occurred. The students successfully identified the linear trend and communicated that greater drop heights yielded greater rebound heights. They noticed the “same ratio” between each group of data points and the “organized scatter” of the data, explaining that there was a relatively constant change between *x* and *y* values. Moreover, students explained that the relationship appeared to be linear because they could draw a line through most of the data points. The 5 samples of student work that Debbie provided were consistent in that students provided written responses to reflect the ideas they had developed during class. Students’ written explanations on their handouts were generally consistent with the high level of thinking and reasoning evident during the observations, yielding an Implementation: Student Work score of 4.

Debbie’s Analysis of the Task. Debbie classified the original, MMR-version of the “Follow the Bouncing Ball” task as a *procedures with connections* task because it allows students to make connections between numerical data for the bouncing of a ball and the mathematical model that is generated through the use of technology. She believes that to be a *doing mathematics* task, “everything has to come from the student,” meaning that students are in full control of the direction of the task: the mathematical question that is addressed, the variables of interest, the data collection procedures, and so on. “At this stage, they’re still kind of following directions,” limiting many of the MMR tasks to *procedures with connections*, though Debbie felt that some tasks contained elements within the *doing mathematics* category. She also stated that *memorization* and *procedures without connections* tasks are more akin to the “traditional or made-up word problems” that are common in school mathematics instruction, differentiating them from most of the MMR tasks.

The way that Debbie planned the task was also at the level of *procedures with connections* in her mind. She felt that the task had the potential to rise to the level of *doing mathematics* but that her students "need a little bit more experience" before they could direct the task on their own. If the task were used later in the year, for example, Debbie explained that her students might not need the same level of guidance and she could structure the task less by removing some of the directions from the handout. She said that, for example, she could have students determine the dependent variable to move the task toward *doing mathematics* rather than directing students to choose rebound height as the handout does. The original MMR version and her own adapted version of the task are definitively at the level of *procedures with connections* to Debbie because she is "asking all the questions" and "leading them [her students] to ask a specific question." However, Debbie also noted that the MMR tasks provide less guidance from one theme to the next throughout the year and move toward *doing mathematics* as a potential goal for students to reach by the end of the year.

The only phase of the task that Debbie designated as *doing mathematics* was the setup phase: as I described previously, she felt that tasks as this level were entirely student-led and teacher-facilitated. Debbie explained that the setup phase was different because "even though I wanted them [her students] to go a particular direction, I think it was coming from them." The students did this by identifying variables of interest and determining if and how they might be measured. Students were also asked to consider the functional relationship between the variables that they had recognized. According to Debbie, her students considered the relationship between drop height and rebound height and analyzed task constraints by using the materials available to determine how they would measure the variables. Students "did what a scientist or mathematician would do in terms of coming up with: What's the question? What are the variables? How are you going to test that?"

By working through the handout and focusing on teacher-generated questions such as, "What is the relationship between drop height and rebound height?" Debbie felt that the implementation of the task was at the level of *procedures with connections*. She was pleased with how her students identified the presence of a linear relationship between the two variables on their own and made the connection that dropping a ball from a greater height led to a greater rebound height. Debbie acknowledged that there were some procedures that she showed them how to do in a "step-by-step" way, such as "setting up the data table." However, she acknowledged that her students made connections between numerical data and graphical representations. She also noticed students making these connections through their dialogue in class: for example, she noted how one student recognized that the regression equation they calculated using technology was linear by the $y = mx + b$ format of the output. She also mentioned that elements of all four TAG levels of cognitive demand may have been evident in the task in some form. For example, plotting points graphically and using software to calculate a regression equation were procedures that Debbie identified, suggesting that the task might include elements of *procedures with* and *procedures without connections*.

Phase 1: Task Selection. Debbie tends to select tasks that are "hands-on" and provide students with opportunities to explore "something that they can touch in some way." These tasks typically prompt students to describe patterns as a whole group, followed by "some form of an experiment where they [her students] can get some sort of an equation," followed by a full-class discussion where students combine their thinking together and formalize their results. She does this for both mathematics and science classes, though she explained that it is sometimes difficult to do with various mathematics topics, such as polynomial functions. "Exploring first" has been Debbie's motivation in both mathematics and science in the past, but the MMR course has encouraged more frequent use of such tasks in her other mathematics courses. She explained,

> With my math classes, I'm noticing that a lot of the stuff I do with MMR is bleeding over into what I had been doing before and that my other classes are becoming more exploratory than what I'd done before.

Some of Debbie's tasks are more structured, prompting students to "follow the directions." She does this to support students who are "timid about trying to do some things for themselves." But many tasks, such as "Follow the Bouncing Ball," Debbie selects so that students will be challenged to "support their reasoning, especially with data."

Emphasis on collecting and analyzing data is an aspect of Debbie's mathematics teaching that may stem from her role as a science teacher. She incorporates as many opportunities for students to work with data as she can because she feels that students "don't have a lot of skills" working with data by the time they reach her classes. Debbie explained that her students tend to be "naïve sometimes in their thinking when it comes to data," though working with data is a skill that is important to "all students, not just the students who are planning to major in a math or STEM career." For example, she wants her students to learn how to make informed decisions about data and to be able to "describe it [data] critically." To help her students develop those skills, Debbie frequently has them "explore" and "describe" data through the tasks she selects. She waits to "put math equations in" until her students have been able to explore some phenomena "to see what is happening." Doing so, according to Debbie, helps her students to make sense of the mathematics and connect mathematical representations (e.g., equations and graphs) back to the original data. "They [her students] need to understand what's happening or the math doesn't really make sense." Specifically, Debbie feels that the "Follow the Bouncing Ball" task provides data that are "strikingly linear" and helps students to clearly visualize a linear relationship between the dependent and independent variables.

Identifying patterns and working with data are important to Debbie, and she has been engaging her students in such activities before she began teaching MMR. In fact, she has been doing this "from day one" of her teaching career. However, she now finds herself "tying it more to a task for a real situation, beyond exploring patterns." Teaching MMR has provided Debbie with frequent opportunities to engage students in working with data to solve genuine, real-world problems because the course has been designed for her to be able to do so. Debbie explained that in some mathematics courses, such as Algebra 2, it is more difficult for her to connect concepts to real-world applications. Teaching the MMR course has given her some "good ideas" to use in her Algebra 2 class, she said. According to Debbie, using such tasks is beneficial for students because they learn skills that are useful later in their lives. For example, being able to describe and interpret data are skills that help students to determine the validity of others' arguments.

Phase 2: Task Planning. Though Debbie chose not to make significant changes to the original MMR version of the task, one concern was being "careful about telling them [her students] too much." She felt as though the original task, along with other handouts, provided more guidance for students and teachers than was necessary. "Sometimes, the handout does a little bit too much thinking for them [the students]," Debbie explained. There were several instances where the "Follow the Bouncing Ball" handout provided written text that Debbie felt she could get her students to come up with through discussion: for example, she explained that she could likely get her students to determine a definition for "rebound height" (the maximum height that a ball reaches after being dropped and bouncing off the ground) and that her students did not need the definition provided for them on the task handout. She said,

> Sometimes I think there's a danger of, you do all the thinking for the student and then they don't have the opportunity to struggle with it a little on their own. So, I mean, that's a skill that we're really trying to build.

Debbie also disliked how the first two problems on the task handout prompt students to think about potential independent and dependent variables, but then the task directs them toward using rebound height as one of them (ideally, the dependent variable, with drop height as the independent variable). However, it is worth noting that Debbie did not modify these aspects of the task.

Thinking ahead to the setup and implementation of the task, Debbie anticipated how her students would engage with the task and the instructional moves she might make as she planned her instruction. Part of this involved considering how her students would collect data: "the more that they are involved in designing the experiment… the more successful they will be," she explained. Throughout her years of teaching, she has noticed that her students are more successful if they are provided sufficient opportunities to talk and write about what they will do to design and carry out an experiment themselves. Debbie also thought that her students might struggle to understand the concept of linear regression; she explained that, after using technology to generate a formula based on their data, her students would need "some practice" and more experience to understand what the formula represented. To support her students' understanding, Debbie considered having her students refer back to the "Starbursts™ Grab" task that they had already done.

As I described in the previous sections, some of Debbie's changes to the task handout involved minor changes to the formatting of the handout. Specifically, some of these adjustments included the materials and the technology that students would use to access the task. In a traditional classroom setting (prior to the COVID-19 pandemic), Debbie would typically have students write what they noticed and wondered about the real-world situation by writing on sticky-notes that they could post on the whiteboard. Instead, she had her students do something similar via Padlet, technology that allows users to post comments in response to a question or prompt. According to Debbie, Padlet benefits students who "might be a little timid at speaking up out loud" and allows students to organize and rearrange their thoughts, something that they would have done with sticky-notes in a typical setting. Another technological change that Debbie made was the use of Data Classroom software, rather than graphing calculators or Excel; Data Classroom allows students to enter and analyze data but does so in a more “intuitive” way, allowing Debbie to focus her instruction more on the mathematics and less on teaching students how to use the technology.

Unlike the other teachers who used “Follow the Bouncing Ball” for the purpose of this study, Ethan and Fred, Debbie planned not to use Part 2 of the task that was described in the MMR lesson plan. Part 2 engages students in an exploration of how the rebound height of a ball decreases exponentially with each successive bounce, an extension of Part 1 where they consider only the first bounce over various drop heights. Though Ethan and Fred mentioned using Part 2 after Part 1 was finished, Debbie explained that it “deviates from the original goals… and seems to go off in another direction.” Because her goal was for students to develop an understanding of linear relationships, she felt that Part 2 “went in another direction and was not adding anything new.” According to Debbie, it shows students that “not all regressions have to be linear,” which could be interesting for students to learn, but was something that could be done later in the course if time allowed. At the present time, however, her goal was that her students expand their knowledge and understanding of linear relationships rather than shifting their focus to other functional models.

Phase 3: Task Setup. Debbie typically sets up tasks through discussions similar to the “What do you notice? What do you wonder?” routine that is emphasized with the MMR course. She especially includes such explorations “towards the beginning of a unit," but it also "depends on where they [her students] are in the development of a concept." The exploratory nature of tasks at the start of a unit allows students to engage with the content more broadly before narrowing the focus to specific concepts later. Debbie began task setup by having her students watch the 1965 Super Ball television commercial, as I described in previous sections; though this was suggested on the MMR lesson plan for the task, Debbie also thought that it helped to increase students' interest and engagement with the task. "They were really interested in that… more than I expected!" she exclaimed, recalling how her students began researching the current price of a Super Ball, whether they were still produced, and what their cost was in 1965. Showing the video helped students to relate to what they were about to explore but also brought forth stronger ideas later during the lesson, Debbie explained. "I think that little video first helped to get them thinking, because I think if I hadn't had something else there first, they might have sat there for a while on that first question on the Padlet." She said that watching the video "caught their [students'] interest" but also helped them to think of things they might not have thought of before.

Debbie used Padlet so that her students could post, comment on, and organize their thoughts as she typically would have prior to the COVID-19 pandemic. During this phase of instruction, Debbie felt that her students were "going in a lot of directions," which she encouraged, but also wanted someone to mention the drop height and rebound height of a ball as variables that could be measured. To guide her students' thinking in this direction without telling them directly, Debbie picked up an elastic ball and started bouncing it, asking students what they observed. It was something that she had not anticipated doing, but thought, “surely they [her students] see what I’m doing.” She reflected, “that helped, I think, in leading them to what I wanted them to do with this activity.” She made the decision in the moment because “a lot of the kids tend to be visual” and she felt that they needed to see the bouncing of the ball to go in the direction that she wanted them to go.

During the discussion of dependent and independent variables, Debbie made several instructional moves to help her students distinguish one from another and recognize whether a variable is dependent or independent. She described how “Those are usually words that they [students] have heard before, but I’m not sure exactly how it’s been used and what they’ve seen,” suggesting that some students struggle with identifying dependent and independent variables. Debbie reflected that her students did not seem to have too many difficulties, but that was why she repeated the question, “Is it influencing the bounce, or it is the bounce?” To help her students identify independent variables, she asked the question, “What is the thing that you’re changing on purpose?” to help students identify that independent variables are often manipulated by the researcher. Debbie wanted her students to recognize that these ideas were important, so she wrote the following on the board as they discussed potential variables of interest: “independent variable: what you change,” and “dependent variable: response.”

Phase 4: Task Implementation. Debbie frequently lets her students explore mathematics and science because she wants them to develop their own understanding; to “see what is happening” on their own. Indeed, this is what Debbie did during the 2 observations of her MMR class. She offered autonomy in how her students could collect data and left them to determine their own meaning for which ball was the “best bouncer” without directly telling how to make such a decision. However, she reflected that there were opportunities with the MMR class where she limited the amount of exploration that her students did for two reasons: (a) as I described earlier, she felt that her students still “needed more experience” before they would be ready to direct themselves entirely and reach the level of *doing mathematics*, and (b) she focused on the question that was emphasized in the MMR materials, guiding her students toward investigating the relationship between drop height and rebound height. Though Debbie acknowledged other ideas that her students may have been more interested in, she wanted them to go in the direction suggested by the lesson plan for the time being. However, she is working toward getting her students exploring fully at the *doing mathematics* level as the year progresses.

Debbie frequently questioned her students throughout the observed lessons. For example, she asked them to explain how their scatterplots provided evidence of the linear relationship they identified. In addition, Debbie typically answered students’ questions with her own questions. She does so because if she answers a question directly, she feels that students will “stop looking at other alternatives” and their thinking process ends with her response. Instead, she wants her students to keep thinking and try to come up with an answer on their own. “In terms of it [the mathematics] sticking, I think it sticks better if you’ve explored that as much as you can yourself,” Debbie explained, recognizing that students will retain the mathematics content if they have done more of the thinking themselves. In our second interview, Debbie recalled an instance where a student had nearly identified the linear pattern in his data. She wanted to provide the name for the type of function, but instead asked the question, “What do you call that?” so that her students would reach the conclusion on their own. After several attempts to explain the pattern, additional prompting, and some wait time, another student eventually used the term “linear” to describe the relationship. In general, Debbie attempts to respond with questions when a student’s initial question addresses an idea that she is “really wanting them to understand.”

Ethan

Ethan has a bachelor’s degree in mathematics, a master’s degree in educational technology, and a principal’s license in addition to his teaching license. The 2020–2021 academic year was his 20th year teaching high school mathematics, all at the same school, and his second year teaching the MMR course. Remotely from home, Ethan teaches a group of 10 students, meeting over Zoom 3 days a week: 35 minutes on Mondays and 60 minutes on Tuesdays and Thursdays. My co-observer and I viewed the implementation of Ethan’s task remotely by joining his Zoom sessions the same way that his students do. We chose to observe on a Tuesday and a Thursday to make the most of our observations, though one student was absent both days. The IQA and TAG classifications for Ethan’s task are presented in Table 9 to provide a summary of the following two sections.

Table 9

IQA Scores and TAG Classifications for Ethan's Task

Score Level	IQA Score	TAG Classification	Mismatch
Phase 1: Selection	4	PwC	No
Phase 2: Planning	4	PwC	No
Phase 3: Setup	4	DM	No
Setup 1: CF	NR	NC	No
Setup 2: MR	2	NC	No
Phase 4: Implementation	3	PwC	No
Implementation: Student Work	NR	NC	No

Analysis of Ethan's Task. Ethan's task was identical to Debbie's at Phase 1: Selection and received an IQA score of 4 for the reasons described previously. Like other teachers, Ethan chose not to modify the "Follow the Bouncing Ball" task aside from reworking the student handout into a series of PowerPoint slides. The written language of the handout remained identical as Ethan simply copied and pasted appropriate text onto each slide and left space for students to type their responses. One minor adjustment was the inclusion of a table for students to record their data. Each row was labeled with units, indicating that students should record their measurements in centimeters. The inclusion of this table lowers the cognitive demand of the task in some regard, but not enough to warrant a decline from an IQA score of 4 to a score of 3 for planning in Phase 2:

Planning. The remainder of the task is identical in terms of the level of work and thinking required of students, including the prompts on the original handout for students to explain their reasoning in writing.

Debbie, Ethan, and Fred set up and implemented the task with some similarities and some differences, despite using the same original task with limited changes for planning. Unfortunately, my co-observer and I were unaware that Ethan's class had already discussed the real-world aspects of the task until after we had observed his instruction on the first day. However, Ethan explained afterward that he and his students watched the 1965 Super Ball video and discussed real-world aspects of the task during their Monday class that week. Therefore, Setup 1: CF was assigned a score of NR because this initial discussion was not observed. However, Ethan led his students in a discussion of the mathematical relationships relevant to the task during the first observation, including potential variables of interest and how they might be measured.

Though the students identified drop height readily, Ethan had to press them repeatedly to recognize rebound height as the other variable to measure. After each student suggested an alternative idea, he asked "What else can we measure?", eventually suggesting "how far it bounces back up" himself. After identifying the two variables, Ethan questioned his students until they arrived at the conclusion that drop height was the independent variable and rebound height was the dependent. Setup 2: MR was assigned a rating of 2 in this instance because, though students provided some ideas and contributed to the discussion, the presence of accountable talk-moves was not consistent enough to warrant a score of 3 (two instances were observed during the 25-min discussion). Phase

3: Setup in its entirety scored as 4; though Ethan suggested one of the variables to include in the preliminary discussion, it was already provided for students on the MMR task handout and the discussion did not reduce the expectations for students when working through the task handout.

The observed implementation of the task declined to a score of 3 on the IQA because Ethan removed some of the challenge and did the thinking *for* his students several times throughout the implementation of the task: for example, he guided students' work when graphing their predictions for drop height and rebound height. He did so by sharing his screen, drawing and labeling the *xy*-plane with the two variables of interest, and asking students what an associated rebound height would be for a given drop height. In essence, he overstructured the task by providing the procedures that students should do rather than allowing them to work through these issues on their own. Ethan also had difficulty obtaining answers from some of his students when he asked them to answer a question or to explain what they had done; the dominance of 2–3 students' verbal responses made it unclear whether the entire class could explain their mathematical thinking and reasoning to the same extent. However, the score of 3 is warranted because students were able to identify the linear relationship between the drop height and rebound height of the various balls that were shown in the MMR data collection videos for the task. After Ethan's questioning about the linear equations associated with each type of ball students used (e.g., golf ball, ping pong ball, and racquetball), they were also able to identify the slope as an indication of the best bouncer. However, there was no conclusive explanation for *why* a greater slope led to a more efficient bouncer.

Because Ethan's students worked in two groups to complete the task and submitted only one assignment per group, I cannot reliably assign an IQA score for Implementation: Student Work in this instance. Therefore, I assigned the code "NR" for this category. However, the sample of work from each group is remarkably different; the mathematical work and writing is stronger for one group than the other (as was the case with Adam, whose students also submitted work together in two groups). Both samples provide numerical data and calculations to answer the prompts on the handout, but the written explanations are more sophisticated for one sample than the other. For instance, when asked to explain what functional model will best represent the relationship between the two variables, one group wrote that a linear model would be best because "The dots on our scatter plot are aligned in a way that a line can be drawn almost completely straight through them," recognizing that their data fit a relatively linear pattern. The other group's explanation, however, focuses on how to correctly input the regression equation into Desmos so that it produces the desired output; this response is not mathematically conceptual in nature.

Ethan's Analysis of the Task. Ethan's TAG classifications for the Bouncing Ball task aligns with Debbie's and, interestingly, for similar reasons. Ethan also classified the task as *procedures with connections* at each phase except for setup, which he rated as *doing mathematics*. He felt that *procedures with connections* was appropriate for Phases 1–2 because students are required to execute procedures (i.e., collecting and graphing data, calculating a regression equation, and so on), but they are expected to come to their own conclusions without the teacher providing a worked-out example. "Even in the way the worksheet's made up, it's like, check with your teacher here and they're not telling you to, you know, tell them the right answer," Ethan explained. Along the same lines as Debbie, he felt that *doing mathematics* tasks were more open-ended and student-led, though his understanding of such tasks seems to be less complete and sophisticated than hers:

> When we have talked about the differences in these tasks with the math modeling group and the professional development, I feel like the *doing math* is more just kind of saying, 'Hey, here's this problem: go!' Like, not having a worksheet.

Like Debbie, Ethan similarly considers that *doing mathematics* tasks begin with a problem situation that the students must explore for themselves. However, the presence of a worksheet does not necessarily suggest that a task is not at the *doing mathematics* level. Such tasks are also more complex than simply setting students free to explore.

Without substantial modifications made to the task, Ethan felt that the task was still at the level of *procedures with connections* at Phase 2: Planning. He noted that students were expected to collect data and represent it in multiple ways, including graphically and using equations. He also reiterated that he was suggesting pathways for students to follow throughout the task without telling them exactly what to do and how to do it.

Ethan explained, however, that the setup of the task was at the level of *doing mathematics* because, "It was very open-ended," and students explored their own ideas freely. He stated that, "The beginning of any good math task is sitting there, saying, okay, what do I know, what do I want to know, and how am I going to figure it out?" Asking students what they noticed and what they wondered and having them identify potential variables of interest reached this level of cognitive demand for Ethan. However, he then felt that moving through the handout and guiding students through the process of thinking, “What’s your *x* variable? What’s your *y* variable? Plot this data…” and so on forced students to focus on procedures rather than allowing them to determine the direction that the task would take. Ethan wanted to deviate from the handout and let his students explore, but felt confined because he was responsible for achieving specific goals:

> I almost wanted to change it up a little bit, and I'm like, no… this is what they [ODE] came up with, I'm going to stick with it. And so, I feel like as it went along, it started losing some of that integrity of being a rich mathematical task, but I also feel like it was almost necessary so that it would push them to get to that point, otherwise they wouldn't have gotten there.

Though it lowered the cognitive demand back to *procedures with connections* for implementation, the structuring of the task was necessary to Ethan because he might not have been able to help students achieve the goals for the task otherwise. However, he also considered that he might have been able to do so in a face-to-face setting because of his ability to move from group to group, provide specific comments and feedback, and better facilitate each group’s engagement.

Phase 1: Task Selection. Ethan explained that the 2020–2021 academic year brought new challenges and introduced new concerns when selecting tasks for instruction. His district shifted from hybrid instruction to entirely remote instruction after the first 3 weeks of the school year, which led Ethan to especially consider tasks that would "translate well to remote learning" and also those that might be interesting for students. He chose not to use the "Barbie Bungee" and "Starbursts™ Grab" tasks, though he thought they were "great activities," because they were not readily suitable for an online learning environment and required the use of hands-on manipulatives. Ethan reflected that his 20 years of teaching have helped him determine "things that are going to work" and things that will not, but remote teaching was an additional challenge that required adjustment. Because he taught the MMR course in the previous year, he had some expectation for which tasks were more easily adaptable and which were not. Ethan recalled the "Follow the Bouncing Ball" task from before, saying "I remember this task from last year. This can translate with the videos, and away we go." He could not take a Super Ball to each student's home, but another MMR teacher had already created videos of various bouncing balls that would assist students with remote data collection.

Ethan also described the importance of tasks that connected mathematics to the real world. Fortunately for him, the MMR course was specifically developed with a focus on mathematical modeling and real-world connections. Ethan explained,

> I feel like a lot of these students that are in my class haven't seen how math affects their daily lives… whether it be in middle school, just the rote algorithms of doing addition, subtraction… Some of these things, they just don't see the importance of whatsoever. And I think, when they can see that connection, it encourages them to do that much more, and to do that much better.

To Ethan, connecting mathematics to the real world and his students' lives helps motivate them to be interested in what they are learning and helps them to be more successful in what they do. He reflected on how he had recently engaged his students with the "Ramp It Up" task (King, 2015) and how for some of his students, the task "hit home" because they either lived with a family member who used a wheelchair or were familiar with someone else who did. "Setting up a wheelchair ramp is pretty real-world for them," he said, and connecting the idea of a ramp to its slope and the associated ADA specifications was meaningful to them. Ethan thought that he and his students had engaging conversations about the ADA specifications for the slope of a ramp and why they were necessary to aide in the use of a wheelchair.

Teaching the MMR course has influenced Ethan to apply a more "task-focused" approach to his AP Statistics and Precalculus classes. He is now using activities from Stats Medic (https://www.statsmedic.com/) and Calc Medic (https://www.calc-medic.com/), which, according to Ethan, encourage the idea of "experience first and formalize later" and a "learn by doing" approach. Ethan explained how the use of these two resources has helped him to translate "hands-on learning" and group projects into his other courses. Teaching MMR encourages him to use mathematical tasks more frequently in his other courses, particularly those emphasizing active, student engagement. Ethan had an "eye-opening moment" teaching MMR when he noticed changes in students who had previously disliked mathematics and not experienced much success with it in the past. After witnessing students experience success with MMR and realizing, "That's doing math," Ethan has shifted toward the use of tasks that promote active engagement.

Phase 2: Task Planning. Ethan described himself as "not the ideal planner" in terms of his regular lesson-planning process. Planning for the long-term is difficult for him because of how drastically plans can change after a single day of instruction. Therefore, Ethan makes a general outline for how he anticipates a task or lesson will go but then flexibly adapts based on what happens in class each day. His need to be flexible was even greater during the 2020–2021 academic year because of the everchanging influence of the COVID-19 pandemic. Prior to each lesson, Ethan considers what he needs to do to get students started, typically by reviewing what they had done in the previous class session. Continuing a task or lesson over multiple days has been difficult because his students do not meet every weekday as they would when learning face-to-face. "Sometimes it's been 24 hours, sometimes it's been 48 hours… and sometimes, it's been 2 weeks," he said, after describing how winter power outages and shortened school weeks had resulted in nearly 2 weeks between classes before our first interview.

Ethan also adjusts his plans when students are struggling, or if many students are absent:

> I'm a big fan of slowing down when kids are struggling, and so I have this planned out for what I think it's going to look like tomorrow, but if we don't get through it all, then I'm going to pick up where I left off.

He would rather take the additional time to work through a task if it means that his students might better understand what they are doing and why, rather than rushing from one topic to the next. Additionally, Ethan adapts his planning when students are absent; with only 10 students in his MMR class and a focus on group work, he must anticipate challenges such as, “What am I going to do with the three kids that weren’t there on Monday?” This requires him to rearrange his groups in a meaningful way in advance so that those who missed class can successfully engage with the next day’s lesson.

For the MMR course, Ethan consults the associated lesson plan for each Context with specific attention to the objectives that are provided. Though he recalled many of the tasks from the previous school year, he made sure to note any changes that ODE made to the tasks for the 2020–2021 year. Because the course is still under development and pilot testing, the materials may potentially change from one year to the next. Ethan explained, “I try and recall what we did last year and see if there’s any changes… they [ODE] have made a lot of improvements and changes for some of these [tasks].” For example, the team at ODE incorporates teacher feedback each year of the pilot, including recommendations, clarifications, and corrections to support teachers’ use of the materials.

After considering the goals for students provided in the MMR course materials, Ethan made additional adjustments to the “Follow the Bouncing Ball” task based on his own objectives for students. For instance, he chose to provide a data table for students to complete with their drop heights and rebound heights. Rather than having students focus on technological aspects of the task using Excel spreadsheets, he wanted to emphasize the mathematics of the task as much as possible in his limited time with students. “I think that they would struggle more with the technology than with the math,” he said, though he added, “Not that I'm trying to make it easy on them, because I think there's something to be said for letting them… figure out, how does equation editor work? Can you make it look nice?" But he expressed that these skills might not be necessary for all students to learn, whereas the mathematics of the task was of greater importance.

Much of Ethan's planning for the "Follow the Bouncing Ball" task resulted from the adjustment to remote teaching. As I described earlier, much of his work involved recreating the task handout as a Google Slides presentation so his students could read, type, and graph on it with their group members. Ethan was cautious of including excessive text on a single slide and also wanted to avoid the activity devolving into "just straight procedures" by simply listing directions for students to follow. In fact, Ethan expressed his desire to use the Desmos Activity Builder instead to transform the handout into an "interactive, group project" that would allow students to plot data and graph equations while simultaneously working through the prompts on the original handout. However, he confessed that he was not yet skilled enough with the technology to do so in a realistic amount of time. Reflecting on the task, Ethan credited the summer MMR PD for familiarizing him with the use of technology such as Google Slides and Google Jamboard that enabled him to engage his students in an online setting. Though he wished that he could have done it another way, he still felt that the MMR PD helped him to do what he might have struggled to do before in his planning.

Phase 3: Task Setup. Ethan thought that the remote teaching environment prohibited students from fully exploring the real-world context of the task. "By being remote, it does really hinder some of that anticipation and the ability to really kind of bring it to the kids," he explained. Ethan showed his students the 1965 Super Ball commercial but acknowledged that his students reacted differently in the remote environment than they had in the previous year, face-to-face. Previously, his students were excited to watch the commercial and were more engaged in it because it was a nonroutine classroom procedure: something for them to be excited about because it was different. Watching the commercial remotely from a computer was, according to Ethan, less engaging because "they're just kind of worn out at this point of watching stuff on their screen." Ethan also suspected that some of his remote students might not have been attentively watching the video in the first place. He reflected that these differing responses to the launch of the task may have led to differences in students' preliminary discussion and engagement. Being able to see and touch various bouncy balls in class during the previous year also evoked student ideas that they did not discuss remotely, including their color and texture.

Similarly, Ethan felt that his students' discussion of mathematical relationships during task setup progressed unlike it might have in a traditional, face-to-face classroom setting. As with Debbie and Fred, Ethan attempted to guide his students toward identifying and exploring the relationship between the drop height and rebound height of a bouncing ball. He thought that his students might have more readily discovered these as interesting, measurable variables to collect data about if they had been in a face-to-face setting. "We eventually got there. I don't know if we would have gotten there faster in person or if it would have been easier to guide them in person, but I would say that was kind of a challenge," Ethan reflected. If he and his students were in a typical classroom setting and they had real, bouncing balls to experiment with, he would have used these materials to his advantage in a similar way that Debbie had done. "If we were in class, I'd be bouncing the ball, I'd be throwing the ball, I'd be dropping the ball… I'd have a meter stick conveniently placed up against the chalkboard.” Ethan felt that this aspect of task setup might have been more efficient, and his students might have noticed different things if he was able to make instructional moves that he could not make as easily in an online learning environment.

Ethan wanted his students to explore during the setup phase of the task. Though his goal was for students to collect data and investigate the relationship between drop height and rebound height, he wanted to guide them toward these ideas without directly telling them what to do. He summarized his thoughts as follows:

> Trying to figure out the best way to get them back to that point without just telling them the answer, I think that's one of the best things that good teachers do well. And sometimes the lack of patience or the frustration level just gets to a point where you're like, ‘it’s drop height and rebound height…’ But I think allowing them to mess around with it and to try some of their own things out is the best way to do it if you're in person.

However, Ethan felt confined in his remote teaching environment and thought that he could have allowed for greater exploration if he and his students were in class face-to-face. Some of this, he attributed to the time that he lost with his students: Ethan explained that, if he had “the full amount of time,” he would have let his students explore other ways to measure how bouncy a ball was, such as how long it stays in the air before coming to a stop. But his remote teaching schedule allotted only 155 min per week with his MMR class, whereas he might have had at least 225 min if they met for 45 min or more each day for 5 days. This noticeable difference in time with students led Ethan to limit his students' freedom to explore so that they could reach the end goal in time.

Phase 4: Task Implementation. As I have described in previous sections, Ethan engaged his students in some form of productive struggle throughout the task. He did not immediately tell students which approach to take but allowed them to direct themselves in some respects. Though it can be a "struggle," especially when students are reluctant to respond, it is something that Ethan finds valuable for their learning. His thought process when implementing a task is consistently, "How do I get them there without telling them the answer?" This is different than how he learned mathematics. He did not do activities as a high school student, and his teachers told him exactly what to do and how to do it. However, he feels that the pedagogy of the MMR course is a "a different way of learning, a different thought process… what we're doing now is better." However, he said that the MMR pedagogy can be difficult for teachers (such as himself) who learned mathematics from a more traditional approach. Even after 20 years of teaching, Ethan reflected that it is challenging to get students to develop their own mathematical understanding.

Ethan frequently prompts his students for evidence of their thinking; though strong explanations were not evident from many of his students over the 2 days, Ethan consistently asked them to make their thinking visible. This is a practice that he has developed over the past several years but does not recall what started it. "It's great to hear somebody explain their answer," he explained, because it encourages students to talk through how they solved a problem, and it also benefits other students who listen in. Ethan believes that asking for student explanations helps him to become a better teacher because it alerts him to students' thinking and possible misconceptions that he might not have anticipated. "It really kind of gives me a little bit more insight into how to get them to the right place, too," he stated. Ethan acknowledged that questioning students and pressing for explanations typically requires more class time, however, it is worth the investment. During task implementation, he asked one student to explain her thinking when attempting to determine which ball was the best bouncer. Though she initially misinterpreted the graphs of students' data, the conversation helped Ethan to understand her thinking guide her toward a clearer understanding of what the graphs represented.

Ethan expressed difficulty in getting his students to communicate and participate in whole-class conversations. "I don't think it was necessarily the math that was the trouble with this… I think it was the participation. Just trying to get the kids to participate appropriately and answer the questions that were being asked," he reflected. Part of the issue, he thought, was again that he and his students did not meet face-to-face in class every day. He felt a weaker relationship between himself and his students because they did not physically see each other 5 days a week as they typically would. Especially when students do not have webcams active, he cannot see them and loses some ability to build stronger rapport with them: "I think by not having the cameras on I think that loses some of the interaction piece. And I definitely miss that because I think this would be a good group."

Another factor that Ethan considered was that he had hardly seen his students over the past 2 weeks due to electrical power outages in the area, combined with already shortened school weeks. However, he did not think that being observed interfered with his students' active participation. In any case, Ethan felt compelled to hold his students accountable by reminding them of their participation grade in the class and frequently calling on those who were reluctant to speak. At one point during the implementation of the task, Ethan told his students, "Come on guys, I need something," expressing the need for them to contribute to the discussion he was attempting to have. Ethan explained later that he does not intend to put students his students "on the spot," but at the same time he expects them to participate and engage with the task at hand. On the other hand, Ethan noted that several students did quite well explaining their thinking. He highlighted the explanations contributed by several of his students throughout the second interview, describing how they contributed to the high cognitive demand of the task.

Fred

Fred has earned a bachelor's degree in music, a master's degree in education and in educational technology, and has taken some coursework toward a PhD in mathematics education. He has taught for more than 25 years in total, with 11 at the school where he currently teaches high school mathematics. Fred was actively involved in ATC programs at Ohio University and has taught the MMR course more than once. He teaches 2 sections of MMR, face-to-face 5 days per week for 43 minutes each day. His MMR students are all high school seniors except for a single junior. My co-observer and I attended one section of the course remotely via Zoom for 4 class periods; task setup occurred during the first day and the final whole-class discussion concluded on the fourth day, with 2 days of student data collection and exploration in-between. We chose to observe all 4 days because we wished to observe the concluding whole-class discussion and had no scheduling conflicts preventing us from doing so. The IQA and TAG classifications for Fred's task are presented in Table 10 to provide context for the following two sections.

Table 10

IQA Scores and TAG Classifications for Fred's Task

Score Level	IQA Score	TAG Classification	Mismatch
Phase 1: Selection	4	PwC	No
Phase 2: Planning	4	PwC	No
Phase 3: Setup	4	PwC	No
Setup 1: CF	2	NC	No
Setup 2: MR	2	NC	No
Phase 4: Implementation	4	PwC	No
Implementation: Student Work	3	NC	No

Analysis of Fred's Task. Fred used the same "Follow the Bouncing Ball" task that both Debbie and Ethan did, and I have already explained that this task received a Phase 1: Selection IQA score of 4. Like other teachers in this study, Fred made few changes to the task when planning aside from reformatting the handout. For example, he included directions for students to watch videos that he had created, demonstrating how to use Excel to calculate a regression equation for a given data set. However, the mathematical content and processes required by students to complete the task remained mostly unchanged. Unlike the other teachers using the same task, Fred added an additional question to the handout after our first interview because it was a question that he had planned to ask in the concluding whole-class discussion anyway: “What does the slope of your best-fit line mean in the context of this data?” As I explain later, Fred felt that the inclusion of this question would raise the cognitive demand of the task, though he still categorized the task as *procedures with connections* before and after the addition.

Fred’s task setup was similar to both Debbie’s and Ethan’s because he also facilitated a whole-class discussion prior to student work time. Similarly, Fred’s students brainstormed attributes of a ball and bouncing a ball that could be observed and measured. Some ideas that they discussed were not mathematical in nature, such as the texture and color of the ball. However, most of the discussion focused on variables of interest and the possible functional relationship between them. Fred elicited ideas from students in both cases as he wrote notes on the board to summarize what they had said. Some examples of variables include the shape, size, and texture of the ball (e.g., if it is smooth or has dimples like a golf ball). The conversation focused on the meaning of a functional relationship: for instance, Fred asked “What are we talking about?” when using the term “functional relationship,” to which students responded that they were interested in x and y variables.

The discussions that occurred during the setup of the task scored as a 2 on both the Setup 1: CF and Setup 2: MR rubrics, but for different reasons. For the CF rubric, the teacher and students must connect ideas about the problem-solving scenario together to warrant a score of 3, which did not happen because students expressed ideas in isolation from one to the next (e.g., Fred asked one student for an idea and then move directly to the next student without linking their ideas together). The MR rubric emphasizes the use of accountable talk moves, which did not occur consistently enough to warrant a score of 3. Overall, the Phase 3: Setup IQA score for this task remained at level 4 because the cognitive demand of the task did not decline from Phase 2: Planning as a result of the aforementioned discussions; Fred solicited ideas from students prior to engaging with the task but did not provide further instruction that would simplify the work and thinking necessary to complete it.

During the implementation of the task, Fred walking around the classroom, monitoring and facilitating students' engagement as they recorded data from instructional videos and worked to answer the questions on the handout. Though Fred intended for students to work individually to avoid close contact (a safety precaution during the COVID-19 pandemic), the students opted to collaborate with each other in small groups from a safe distance. Fred allowed students to work in groups but requested that each student submit their own assignment when they were finished.

As stated previously, Days 2 and 3 of the task involved student data collection and work to complete the task handout; unfortunately, the remote observations necessary for the present study made it difficult to observe and analyze student-to-student and teacher-to-student interactions that might have been observable if we had been present in the room; however, the concluding whole-class discussion on Day 4 spoke to the mathematical rigor of the task throughout the 4 days. To answer the question, "What does the slope of your best-fit line mean in the context of this data?" students provided conclusive statements such as, "For every 10 cm on the drop height, the rebound height increases by 6 cm." This was an interpretation of the slope 0.6 in the regression model $y = 0.6x + 5$. This required some scaffolding from the teacher, who asked questions to help his students reconsider 0.6 as 6/10, a change in y (the rebound height) divided by a change in x (the drop height). However, students frequently provided such evidence of their thinking and reasoning, resulting in a Phase 4: Implementation IQA score of 4.

Fred submitted 12 samples of student work associated with the "Follow the Bouncing Ball" task and I chose 6 to analyze using the IQA rubrics based on the criteria outlined in Chapter 3. The Implementation: Student Work score assigned to the collection of six samples is 3 because, holistically, the samples provide evidence that students made conjectures (predictions about how the drop height influences the rebound height), identified patterns in their data (e.g., increases in drop height led to increases in rebound height), and used multiple representations (numerical data, graphs, and algebraic equations). For example, one student made this connection clear by writing: "Every time the original [bounce] height increases by 10 cm the rebound height increases 6 cm." Students were also able to successfully calculate a rebound height based on a given drop height and even work backwards to calculate an initial drop height from a given rebound height using their linear regression equations. However, the samples of student work generally lacked solid explanations that warrant a score of 4 on the IQA rubrics.

Fred's Analysis of the Task. Fred classified the original, MMR-version of the task as a *procedures with connections* task, emphasizing that the task requires students to collect data, record it in a spreadsheet, and analyze it after having technology calculate a regression equation. He cited the statement, "Although general procedures may be followed… they cannot be followed mindlessly" in the TAG, suggesting that students must think about and understand what they are doing throughout their engagement with the task rather than simply doing what they have been told to do by the teacher. Fred also quoted that the task "requires some degree of cognitive effort" for a similar reason: his students would have to think and reason their way through the task because it was not identical to something they had seen or done before.

When asked about the planning phase of the task, Fred addressed some of the technical changes he had made, such as rewording questions and directions to accommodate for video data collection (instead of actually bouncing various balls) and the use of Excel spreadsheets (the original task suggests that students will use a graphing calculator instead). Fred reflected that these modifications changed the "delivery method" but not necessarily the mathematics involved in the task. Through our conversation, he explained that he planned to ask students the question, "What does the slope mean in this context?" even though it was not explicitly written in the task handout, something that he "often along the way" asks when implementing this task. In fact, Fred decided to make this change to the task handout after discussing it during the first interview. He said, "If I ask that question, then I think that they [the students] are doing mathematics" to some degree. However, it was not enough to raise the cognitive demand of the task entirely out of the *procedures with connections* category: "I think it still stays in the… procedures with connections," he concluded.

Fred categorized the task as *procedures with connections* in Phase 3: Setup as well. He stated, "I don't think it dropped down to ever being *procedures without connections*" because students did more than simply "regurgitating a process" without understanding what they were doing and why. Fred attributed some of the high cognitive demand to the suggestions that students gave during their preliminary discussions, including aspects that would influence the bounce and those that would not. However, he also thought that his students' ability to determine independent and dependent variables was still "rudimentary" after their work with the task. Fred reflected that he "guided them to the fact that bounce height was the independent variable" and wondered if they would be able to identify independent and dependent variables if asked to do so on another task. He used the word "rudimentary" because he thought that his students might not have strong reasoning for labeling each variable as either independent or dependent, and that they might just assign them randomly.

Phase 4: Implementation, implementation, was also at the level of *procedures with connections* in Fred's view. Though he needed to press students frequently to obtain the responses he was looking for, he also stated that "they know what the procedure is: it's taking the values and plugging them into the equation they have found." Fred explained one example where a student asked, "Which variable does this go in for?" To support the student without telling him what to do, Fred asked "What are the meanings of your variables in your equation?" to get the student to think about what each variable represents in the real-world context. In terms of the cognitive demand of the task and the

TAG, Fred explained that "It's still… procedures with connections because it's connected to the data which they collected."

A more traditional word problem where the equation was provided for students and they simply calculated rebound heights for various drop heights, according to Fred, would be the *procedures without connections* version of what he had his students do with the task. If the instructions for the task were, "Use this equation to answer the following questions… then it would be, I think, *procedures without connections*," he said. Lastly, Fred concluded that "giving it [a task] a context" does not necessarily mean that students are making connections. What set the Bouncing Balls task at the level of *procedures with connections* was "work that they [the students] have done to determine the equation on their own or with the help of technology." Fred explained that students made connections by applying the equations that they had determined after conducting the experiment and collecting data.

Phase 1: Task Selection. Like Debbie and Ethan, this was not the first time Fred has used the "Follow the Bouncing Ball" task. But Fred has implemented a similar version of this task numerous times even before his involvement with the MMR course. He explained that the original idea for the task stemmed from a journal article in NCTM's *Mathematics Teacher* "a long time ago," though he did not recall the specific article title or publication year. Thus, Fred was very confident and "comfortable" selecting and adjusting this task for the 2020–2021 academic year because of his vast experience and familiarity with it. He had taught using his own, modified version of the original NCTM task for many years, now using the MMR version of the task since his involvement with the pilot course. The inclusion of the 1965 original Super Ball commercial was an aspect of the task that was included by ODE to “perhaps drive a question” and peak students’ interest. However, other aspects of the task were quite similar to the version Fred had been using prior to MMR. Fred has “a lot of experience doing it [the task]” and acknowledged that he perhaps had more flexibility with using and adapting it than some of the other MMR teachers might.

Fred chooses tasks that engage his students through exploration. This is not only true for his MMR course, which is inherently built around active, student-centered pedagogy, but also in more traditional high school mathematics courses such as Algebra 2 and Precalculus. Fred sometimes lectures in these courses, but also incorporates tasks that engage students in developing deep mathematical understanding. For example, when teaching his Algebra 2 students about parabolas, Fred had them launch objects in the classroom and plot points along their trajectories. Then, he and his students graphed the points and determined quadratic equations for each using Desmos, allowing them to discuss similarities and differences based on the launch angle and the data for each object. In Precalculus, Fred’s students explored algebraic and graphical representations for various rational functions, again using Desmos. By prompting students to describe what they noticed about the equations and their associated graphs, Fred supported his students in making connections to what they had learned about polynomial functions and extending their prior knowledge appropriately in a new context. Fred describes himself as “not your traditional Honors Algebra 2 instructor, and even Precalculus” because he incorporates numerous investigative activities rather than lecturing only.

To Fred, the mathematics of the “Follow the Bouncing Ball” task is important for students to learn. He wants his students to learn more about linear functions, as is the goal of the task; however, he stressed that the task was beneficial because of its potential to get students thinking, “What does this mean?” Through their engagement with the task, Fred hoped that his students would make connections that they might not have made in their previous mathematics courses. “These are seniors who have taken all the courses that may not have had connections to other math they take,” he said. They had already learned about slopes, *y*-intercepts, and how to write the equation of a line, given two points. Fred’s focus for the task was not on “symbolic manipulation,” but instead for students to develop deeper understanding and make new connections to the mathematics they had already learned.

Phase 2: Task Planning. Fred’s planning for the task included anticipating how he thought the lessons might go and considering the instructional decisions he might make throughout. He initially planned for his students to work individually to encourage safe social-distancing amidst the COVID-19 pandemic. He also thought about aspects of the task that might “need some discussion for small groups and perhaps the whole class,” such as the meaning of a “functional relationship,” because one of his goals for students was to develop a strong understanding for how drop height and rebound height (among other variables) are related. Fred anticipated where he thought his students might struggle: one such instance, he thought, was when students reached the questions on the handout focusing on correlation coefficients. Because his students were less familiar with this concept, Fred acknowledged that he would likely do “some type of full-class instruction on that." Students' use of technology was another concern; Fred had created several instructional videos explaining how to enter data and perform linear regression using Excel spreadsheets, for example. "I have to think about where I'm going to give this instructional material," Fred pondered aloud. This was an adaptation that he made for the task that was not in the original MMR handout, so he needed to consider where it would fit in the activity and how he would support students in using the technology efficiently.

Unlike Debbie and Ethan, Fred made it clear that one of his goals for the task was for all his students to become "well-versed" in the use of spreadsheets. Fred appreciates the value of Excel as a tool to record data, create mathematical formulas, and generate models (such as best-fit lines) to represent data. Therefore, he was willing to spend some extra time teaching his students how to use spreadsheets, especially with this task, because it was one of their first major uses of this technology for the year. He had tried to use spreadsheets occasionally during their first quarter, when all his students attended class remotely, but explained how "it wasn't working out" yet.

Another of Fred's goals for the task was motivating students to think about the meaning of the linear models they found and how they related back to the real-world context. During our pre-observation interview, he explained that he planned to ask students "What does your slope mean in this context?" Somewhat to his surprise, he then noticed that this question was not posed on the MMR handout for the task. As such, he revised the handout a second time and included this question explicitly so that his students would be prompted to make the desired connection. Similarly, Fred added a question prompting students to interpret the *y*-intercept for their models so that they might recognize what it meant to have an unrealistic, nonzero value.

As I mentioned in the previous paragraphs, Fred created several instructional videos to assist his students with various aspects of the task. He created one set of videos to simplify data collection so that students could do so safely during the COVID-19 pandemic. These videos depicted the drops and bounces of various types of balls in slow-motion, the same videos that Ethan described using for the task. In a typical school year, Fred might have taken his students outside of their classroom and into a nearby hallway to collect and record data by hand. Instead, he modified the printed text on the task handout to reflect the new remote data collection procedures. Fred created another set of videos instructing students on how to use Excel spreadsheets, as I described previously. He chose to use videos to reduce the amount of time he spent directing the whole class, therefore increasing the amount of time that students could work on their own. Both changes also benefited the few of Fred's students who elected to continue with remote instruction because they could watch the videos on their own time at their own pace from home.

Phase 3: Task Setup. Fred focused primarily on task implementation during his post-observation interview. However, he commented on students' preliminary discussion of variables that might influence the bounce of a ball and how they might be related. "I think it [the discussion] went pretty well," he stated, explaining how students' ideas were "right on" in terms of what would affect the bounce and what would not. For example, his students suggested that the shape (e.g., perfectly spherical or covered with dimples like a golf ball) and the texture (the material composure) were features of a ball that might affect the bounce, whereas the color and the smell of a ball were unlikely to have a significant impact. After their discussion about independent and dependent variables, Fred felt that his students' ability to determine one from the other was still "rudimentary." He said that they "need more work with how that works out," suggesting that they might need more experience to help improve their knowledge. To support his students in developing this understanding in the future, Fred explained that "The biggest thing is to ask them, in some regards, what do they have control of?" Though he acknowledged that "It sort of doesn't matter how you label the axes" in a graph of dependent versus independent variable, he wanted his students to understand that the former is typically set along the vertical axis and the latter along the horizontal.

Phase 4: Task Implementation. Rather than immediately guiding students toward the use of drop height and rebound height as the independent and dependent variables, Fred initially allowed them to consider any variables of their choice to explore. "I wasn't trying to steer them towards any particular way at that point," he explained. He wanted to maintain the open-ended nature of the task until the third problem on the task handout, indicating for students to consider rebound height specifically. In past years, he has even allowed students to consider various independent variables, such as the time it takes a ball to reach the ground. This required more time to implement the task than Fred felt that he had available in the present year, so he chose to limit students' ability to explore more so than he had done in previous years. Despite the apparent pressure for time, he also allowed his students to make and correct some of their own mistakes throughout the implementation of the task. For example, some students had visually identified an incorrect *y*-intercept for their graphs because of how Excel had automatically formatted the axes. Fred "saw that… and didn't say anything" because he wanted them to correct themselves after noticing a different *y*-axis in the regression equation that Excel calculated for them.

During the final whole-class discussion, Fred provided more support and guidance to assist his students in making connections. One instance occurred when the class discussed their linear regression models and what the *y*-intercept of their equations meant in the real-world context. Based on their data, many students found that their *y*-intercepts were nonzero, meaning that a ball would bounce (or fall through the floor) even when the initial drop height was 0 cm, which was unrealistic. Fred's students "were just really looking puzzled" when he first asked the question, and some responded that the *y*-intercept was "where it crosses the *y*-axis."

Though he wanted his students to provide a real-world description, he used what they "gave" him and sketched the *xy*-plane on the board and graphed one of his students' equations as an example. "I drew the axes," Fred elaborated, "because I wanted to… focus their attention on what the equation was saying, then try what that meant." This led one student to state that the *y*-intercept was "where the graph starts," prompting Fred to ask, "What do you mean by 'starts?' What is the drop height there?" By asking the question about the drop height at the *y*-intercept, his students were able to identify that the drop height was 0 cm and that the ball "didn't drop from anywhere." Fred describes this as “just-in-time help,” providing assistance in the form of a statement or question that might prompt his students to come up with the next idea.

Fred frequently questioned his students throughout their engagement with the task. Some of his verbal questions were planned, such as what the *y*-intercept meant in terms of the real-world context. Others, he formulated and asked in the moment to extend students’ thinking further: one such example was when he and his students were discussing what the slope of their regression equations meant in the real-world context. A follow-up question that Fred spontaneously asked was, “Can you have a slope of 10?” He asked this question to help his students realize that a slope of 10 meant a change in rebound height of 10 cm for every 1 cm change in drop height, which is impossible because of gravity.

Another form of questioning that Fred uses is to ask a question when students provide a correct response. He elaborated on this move, saying,

> I often do that… especially if they give a correct answer. I actually ask them, “Are you sure?” I’m trying to deliberately break the mold. Teachers have usually asked “Are you sure?” when the student’s got it wrong… and sometimes they'll change their mind, I think because their experience tells them that it must be wrong.

Fred asked this type of question during the first observation when a student voiced how the color of a ball would not affect its bounce. By questioning students, even after providing a correct or logical response, Fred prompts his students to “second guess just coming up with a ‘yes’ or ‘no.’” He wants them to realize that that they are not necessarily wrong simply because he asks them a question; sometimes he just asks to keep them thinking or so that they might think deeper.

Gwen

Gwen holds a bachelor's degree in human development and family studies and a master's degree in education with a teaching license. She has been teaching for 8 years, all at the same school. The 2020–2021 was her second year teaching the MMR course; she has a single section of only 4 students. Gwen has taught all four of these students before and they have typically "struggled" learning mathematics, though they are also "familiar" with her teaching style and expectations. Gwen taught face-to-face 3 days each week: Mondays for 25 minutes and Wednesdays and Fridays for 75 minutes each. Both observers attended Gwen's MMR class remotely using Zoom and observed one pair of consecutive Wednesday-Friday meetings to optimize the amount of time we would have with the class. The IQA and TAG classifications for Gwen's task are presented in Table 11 to provide context for the following two sections.

Analysis of Gwen's Task. Gwen's task was provided with the MMR course materials but originated as an NCTM journal article titled "Ramp It Up" (King, 2015). The brief article introduces the idea of slope in the context of ramps, with special attention to the specifications designated by the ADA. Afterward, King poses seven problems related to lengths, heights, and slopes of ramps that teachers can share with their students to solve. Within the sequence of tasks in the MMR course, this task is intended for students to solve immediately prior to the "Discovering Slope" task that Beth used. The Phase 1: Selection IQA score for the task is 3 because students engage in problem-solving and are not provided with a direct algorithm for solving each of the problems. Students also may use multiple strategies and representations, including numerical representations (e.g., the ratio 1:12 referring to a ramp with a length that is 12 times its height) and pictorial representations of ramps as drawings or diagrams. However, the task does not explicitly prompt students to provide evidence of their thinking and reasoning. The Phase 2: Planning IQA score for the task is also 3 because Gwen made no modifications to the task aside from excluding two problems on the handout; the remaining problems still warrant an IQA score of 3.

Table 11

IQA Scores and TAG Classifications for Gwen's Task

Score Level	IQA Score	TAG Classification	Mismatch
Phase 1: Selection	3	PwC	No
Phase 2: Planning	3	PwC	No
Phase 3: Setup	4	PwC	No
Setup 1: CF	3	NC	No
Setup 2: MR	3	NC	No
Phase 4: Implementation	2	PwC	Yes
Implementation: Student Work	2	NC	No

Task setup occurred as suggested in the MMR lesson plan: Gwen began by taking her students to the location of a ramp in their school building, where she engaged her students by asking the questions, "What do you notice?" and "What do you wonder?" about the ramp itself. Then a short debrief occurred, during which students shared what they had noticed and wondered about the ramp, leading to a conversation about the incline of the ramp, its "steepness," and ways that it could be measured. Afterward, students collected various measurement data and then returned to their classroom. Gwen led another discussion on ramps, asking questions such as "Where have you seen ramps before?" and "What makes a good ramp?" Students provided brief responses after Gwen prompted them to talk to a classmate, such as the ramp's location and its width to allow more than one student up or down at a time. Finally, the students took turns reading the narrative of the "Ramp It Up" (King, 2015) article before Gwen allowed them to work in pairs and complete the problems on the handout.

Setup 1: CF scored a 3 because Gwen and her students occasionally linked ideas together (e.g., connecting shared experiences involving ramps in the auditorium and in other school locations) and the students participated actively in the discussion. Setup 2: MR also scored a 3 because of the consistent use of accountable talk moves when discussing students' measurements for the ramp and the diagrams they drew to illustrate their measurements. For example, when it was evident that students measured the "length" of the ramp two different ways (one using the horizontal length across the floor and the other using the oblique distance from the beginning to the end of the ramp), Gwen asked her students to identify "Which is a better representation of length?" The students suggested that the oblique length way better and she pressed further by asking them to explain why they felt this representation was better, though they did not clearly provide mathematical reasoning for their choice. The use of accountable talk moves was consistent throughout the setup of the task but lacked a single “strong” move to warrant a score of 4.

The overall Phase 3: Setup score for task setup raised to a 4, however, because Gwen explicitly set the expectation for students to “verbalize” how they approached each problem, communicating her desire for students to provide evidence of their thinking and reasoning verbally. Setting this expectation for students’ engagement with the task enhanced the rigor of the task above what was communicated in print on the task handout.

Task implementation, however, failed to meet the same rigor as was evident in task setup. In fact, I scored Phase 4: Implementation as a 2 on the IQA, below even the original score of 3 in Phases 1–2. The students appeared to be lost and confused when initially examining the 5 problems on the handout, and they required persistent intervention from Gwen throughout the observed lessons. As a starting point, she suggested that her students draw diagrams to illustrate the given information and to determine what information was needed to solve each problem. Students also required additional prompting to understand the meaning of ratios such as 1:12, 1:16, and 1:20 that were discussed in the task handout.

After guiding students toward the process of using similar triangles to identify the missing side lengths for various ramps, the focus of the task shifted to solving equations such as $1/16 = 15/x$ and the meaning for doing so seemed to be lost in the procedures. Gwen frequently prompted her students to explain their thinking, but they provided responses indicating only the procedures they used to solve their equations. For example, Gwen asked one pair of students to explain how they determined that the length of a 15 in high ramp with a rise-run ratio of 1:16 was 240 in. The students simply referred back to their equation $1/16 = 15/x$, where $x = 240$, using the procedure of "cross-multiplying and dividing" as their justification. Not only did the students fail to explain how they set up such equation and how they solved it, but the answer of 240 in describes the horizontal length of the ramp, not its oblique length (the intent of the question). This misinterpretation was not discussed and the actual answer to the problem was not sought.

The Implementation: Student Work score for the 4 samples of student work similarly scored a 2 because students' written work displayed only their diagrams and the calculations they made to find missing side lengths. Students documented some of what they noticed and wondered about the ramp they observed in their school, but their responses lacked substance; for instance, one student simply wrote that she noticed the ramp was "going downhill" and did not provide details or explain any mathematical ideas in depth.

Gwen's Analysis of the Task. Gwen categorized the "Ramp It Up" (King, 2015) task as *procedures with connections* holistically but identified aspects that were either at the level of *procedures without connections* or *doing mathematics*. According to Gwen, the cognitive demand of the task "averaged" around the *procedures with connections* level for this reason. For example, some of the problems require only straight-forward calculations for the slope of a ramp given its “rise” and its “run.” However, other problems require sophisticated thinking on the part of the student according to Gwen. She indicated that several bullet points in the TAG fit the task, including that the task suggests pathways to follow with connections to conceptual ideas and that the task requires some degree of cognitive effort, but general procedures cannot be followed mindlessly (in the *procedures with connections* domain). However, the task also requires students to access relevant prior knowledge about slope and requires students to actively examine task constraints (in the *doing mathematics* domain).

Phase 3: Setup, according to Gwen, remained at the level of *procedures with connections*. Initially, she expressed that her intention was to maintain the cognitive demand of the task. Rather than speaking in terms of the thinking and reasoning that the *students did*, she initially explained her perspective of the cognitive demand of the task in terms of how *she intended* for the preliminary discussion of task to unfold:

> I think I was trying to maintain it [the cognitive demand], perhaps with the students it maybe went up a little bit because they didn't quite understand. But I think that I maintained it. I'd be curious to know what you think. But I definitely think I maintained it, that was my intention.

Contrary to the intention of the TAG, Gwen expressed that the cognitive demand of the task “went up a little bit” *because* students failed to understand what they were doing at times, rather than declining in this instance. It seems she does not understand that the teacher’s responses to students’ difficulty and the way in which students engage with the task influence its cognitive demand. Gwen also explained that the setup of the task required “some degree of cognitive effort” because she “wasn’t just handholding them [her students] and giving them information.” Indeed, she pressed her students to contribute and provide their own ideas to the preliminary discussion.

Though Gwen identified student difficulties and shortcomings throughout task implementation, she still categorized implementation as *doing mathematics* “with the push” from her. She attempted to help her students visualize the task in various ways by suggesting that they draw pictures and diagrams, which students “really struggled with.” According to Gwen, her students were “trying to regurgitate information” that they had recently learned in their Chemistry class for how to calculate slope by using algorithms rather than attempting to conceptualize what they were doing and why. Moreover, Gwen acknowledged that her students “didn’t know or understand why” they could set up and solve a proportion involving similar triangles; they simply applied a previously learned algorithm in a familiar situation. Gwen attributed students’ lack of understanding when applying procedures to their prior instruction with another teacher:

> Their junior high math teacher… dealing with similar triangles, taught them proportions, taught them to cross multiply and divide. Even though both the Algebra 1 teacher and myself tried to unteach that, somehow it still stuck in there. So, it's just ingrained in them from hearing it over and over…they follow these learned habits and learned sayings.

She even stated that her students were “kind of alluding to the connection between similar triangles and proportions… but didn’t really have the connection.” However, she concluded that her students still maintained the level of *procedures with connections* because the task required “some degree of cognitive effort. Even though it may not have been through work of their own… I was pushing them to get to that point.”

Phase 1: Task Selection. Gwen taught using mathematical tasks and activities frequently even before she became involved with the MMR course. She “rarely” provides direct instruction in her Algebra 2 and Precalculus courses and feels that the pedagogy for the MMR course closely aligns with what she had been doing already. Through her experience as a paraprofessional and then as a starting teacher, she has witnessed many students fail to learn and succeed through lecturing and other teacher-centered approaches. Her experience has also taught her that students learn more through engagement and effective teacher questioning. Specifically, she emphasizes the use of “low floor, high ceiling” tasks that are “approachable by all students” by providing “multiple entry points.” Such tasks, according to Gwen, “engage *all* students, not just the ones who ‘get it’ already.” She selects such tasks because she wants to engage students with limited background knowledge, those who might not have experienced success with mathematics in the past. “Not all students have high mathematical self-efficacy, or the background knowledge to approach certain classroom activities, tasks, or lessons,” Gwen explained. “For those students especially, tasks should make them feel comfortable enough to start.” By selecting tasks that are both accessible and engaging to her students, can then scaffold and guide them to the “high ceiling” of deep mathematical understanding.

The “Ramp It Up” task was both engaging and accessible, according to Gwen. Another desirable aspect of the task was that it “connects math content to a real-world problem," encouraging students to explore slope in the context of ramps. The task "pulls students in from the beginning, making connections to something that they see on a daily basis," Gwen explained. She selects many tasks with real-world components to address the frequent student question, "Where will I use this?" Though her students may never need to calculate slope in the future, being able to "approach real-world problems and think critically through them" is a valuable skill that students can learn through the "Ramp It Up" (King, 2015) task. Gwen explained that her students' definition of "doing math" was "plug and chug," meaning that they viewed mathematics as simply regurgitating algorithms and performing routine calculations. By relating mathematics to the real world, Gwen hopes to also shift their beliefs about mathematics and help them realize how important mathematics is in daily life.

Gwen generally selects tasks that are "connected to [mathematics] standards, either past, present, or future." She did not specify which standards the task addressed specifically, but she communicated that the content aligned with her goals for students and what she wanted them to learn. Gwen would use "Ramp It Up" (King, 2015) in other courses than MMR because it "forces students to analyze and not simply regurgitate information." She did not want to simply demonstrate the formula and procedures for calculating slope; she wanted her students to have to think through the process and make sense of information that was given to them. The task provides students with problems to solve but does not specify the approach that students must take; in this sense, it "tells them what to do, but doesn't tell them what to do," according to Gwen. She also appreciates that the task provides students with opportunities to explore multiple representations, including scale models, formulas, and numerical measurements.

Phase 2: Task Planning. Gwen plans her instruction with a focus on the objectives and goals that she has for her students but allows for some flexibility as well. She typically plans an entire unit in advance, "roughly 2–4 weeks, depending on the course or the content studied." From there, she chooses tasks and activities that she thinks will help her students to meet the goals for the unit. However, she also is willing to adjust her plans "on a daily basis" depending on how her students interact on a given day. With this flexibility, she always keeps her "end goal and outcomes" in mind. With the "Ramp It Up" (King, 2015) task, Gwen's goal for students was the following: "Can they interpret slope and use right triangles to solve problems?" Moreover, her intention is "to train her students how to think, not just how to plug numbers into a formula or equation." With these goals in mind, Gwen chose not to significantly modify the task handout but decided to eliminate the first page of the "Discovering Slope" student handout (recall that this was the task used by Beth). According to Gwen, the first page "implies that direct instruction will occur" and would rather her students ask questions about the information on that page rather than providing it immediately herself.

Planning the implementation of the "Ramp It Up" task encouraged Gwen to consider her students' prior knowledge and anticipate both how they would engage with the task and how she might respond. She acknowledged that her students had recently studied slope in their Chemistry class but felt that their instruction was likely limited to routinized procedures. Therefore, she anticipated that their prior knowledge of slope would lack connections and deep understanding, leading to some amount of difficulty with the task. Gwen explained that her students were not accustomed to "pushing their thinking," leading them to struggle "interpreting the questions" and identifying "how it [slope] applies to this task." Moreover, she expected that they would have "little to no recollection of the Pythagorean Theorem or similar triangles." Though she appeared to have low expectations in terms of her students' prior knowledge, she felt that she could help her students be successful with the task through "continued repeated questioning." Gwen felt confident in her ability to provide the necessary scaffolding to support their engagement and high-level thinking throughout their work on the task.

Phase 3: Task Setup. Gwen intentionally engaged her students in a discussion of ramps during task setup. "I think it's important to know internally what the students know… before you dive into something like that," she explained. She did not expect them to all be familiar with the real-world context prior to their work on the task. However, Gwen reflected that her students did not fully appreciate or connect with the real-world application of ramps as she hoped they would. "I honestly don't think they fully still understand the real-world connection, at least to a level that I would hope that they would at this point. But it at least got their gears turning." Gwen recalled students' responses in their preliminary conversation, such as how the ramp they examined was "smooth," "had rails," and that ramps should be "wide enough for multiple people to go down" at a time. However, her students did not provide the "genuine, deep responses" that Gwen sought. She was given the impression that the real-world connection to ramps was not interesting and relevant to this group of students: "My understanding is none of the four girls have an actual connection to ramps. They don't have family members who need to use [them], houses that have to have a ramp." This, according to Gwen, led her students to disengage with the task to some extent in both setup and implementation.

As Gwen presented her students with the task handout and had them take turns reading the accompanying passage, I noticed that they did not understand the meaning of ratios such as 1:12, 1:16, and 1:20. As one student read over the text "1:20," for example, she hesitated and said the phrase, "one-twenty" rather than using language to indicate an understanding of 1:20 as a proportion, such as "one-to-twenty" or "one is to twenty." During implementation, I also noticed that Gwen spent some time helping each of the two groups interpret what these representations for ratios meant. Therefore, I asked her what she thought might have happened if she had led her students in a discussion of these relationships prior to their engagement with the task to help develop a common understanding before they started. Gwen explained that the implementation "would have been worse" because her students might have "shut off" and not thought about it on their own. Doing so, according to Gwen, also would have established a standard that she was going to provide all the necessary information for her students, which she wanted to avoid. "I'm sure you saw it, they shut down when they don't immediately understand it [a task] because they are very used to being given a procedure or algorithm and just repeating," she answered. Alternatively, taking the time to help students develop their own understanding for ratios written as a:b might have supported them to be more productive during implementation.

The distinguishing feature of Gwen's task setup that led to an IQA score of 4 (an increase from the Phase 3: Setup score of 3) was that she provided the verbal expectation for her students to make their work and thinking visible. She specifically prompted her students, "Verbalize how you approach each problem" before she instructed them to start working in their pairs. By doing so, she introduced an element to the instructions for the task that was not evident in the handout. Gwen typically sets this expectation for her MMR students because it enhances her understanding of their thinking processes. "I can't fully grasp what they're thinking just by looking at what they have on paper… it's more than being able to reproduce something, you have to be able to explain what's going on as well." However, Gwen recognized this as something that her students have generally "struggled" with. Despite this being her expectation for students in task setup, the implementation of the task failed to meet this standard.

Phase 4: Task Implementation. Gwen tries to promote productive struggle through the use of engaging tasks and avoids direct instruction. She described how, especially early in the year, she found it important to "set the tone" for the MMR course by reminding her students that they would not learn through direct instruction. Instead, Gwen's goal for her MMR students was to learn how to learn and do mathematics in a new way. During the implementation of the "Ramp It Up" (King, 2015) task, Gwen was "disappointed" that two of her students (one out of each pair) "took over" most of the thinking and the other pair seemed to shut down. According to Gwen, several other factors worked against her students: she noticed that one of them appeared very lost, but "wouldn't verbalize that she didn't understand," perhaps because my co-observer and I were watching her from Zoom. Her students had also played an important volleyball game the night before, leading them to be exhausted and less "chatty" than they typically were in class.

In addition, Gwen thought that her students might have been afraid to make mistakes publicly in front of each other; though they were all friends, they might have been "hesitant to share information and be wrong" in front of each other and in front of us as observers. Throughout various interactions with her students, Gwen gathered that they had negative experiences with mathematics in the past, including "teachers who have… alluded to the fact that they're slower than their class." This may have resulted from either being put into "lower" groups based on their performance or receiving low grades in other mathematics courses. Gwen explained that these negative experiences may have influenced her students' reluctance to communicate their misunderstandings, even though she has emphasized that it is "completely okay" to make mistakes in her class.

Gwen asked her students many questions during the observed lessons to understand their thinking and to help guide them from one idea to the next. "It just goes back to forcing them to think," she elaborated. "I don't want to be the one that gives them all the information." Instead, she asked scaffolding questions so that her students would come up with the mathematical ideas on their own. As her students engaged with the task, Gwen listened for things that would prompt her to ask a specific question. "I'm constantly trying to listen for what they [the students] understand." Despite their limited success with the task, Gwen expressed some positivity in how her students responded to her questioning. “They didn’t completely shut down, so that’s positive… I think this is the first task they’ve had where they had no idea where to start. So, I guess in that regard, they responded really well,” though she also thought that the quieter students could have done more to communicate their thinking. Gwen stated that she would not change this aspect of her instruction because she wants her students to become more comfortable with being questioned explaining their thinking verbally.

In addition to questioning her students persistently, Gwen provided additional support in the form of suggestions. One instance of this involved her response to students’ initial frustration getting started with the task. Specifically, her students did not understand the meaning of the ratios 1:12, 1:16, and 1:20 when reading them on the task handout and therefore did not visualize what they represented. To assist her students, Gwen asked one of them to draw 3 right triangles (representing the side-view of three ramps) on the smart board and to label their side lengths based on each ratio. Her students successfully labeled each with a height of 1 unit and horizontal lengths of 12, 16, and 20, allowing them to better understand how the three ratios might be comparable to each other. This led back to the question of, “Which is the shortest ramp?” and enabled Gwen’s students to progress further through the task. During our second interview, she later explained that she “had to intervene” at this point because her students would “completely get lost” otherwise.

Henry

Henry has a bachelor’s degree and a master’s degree in education and has completed some graduate-level mathematics coursework. He has taught high school mathematics for 34 years with 21 at his current school. The 2020–2021 academic year was not Henry's first year teaching MMR. Henry teaches 2 sections of MMR that meet both face-to-face and remotely, though he is in the classroom every day; one group meets face-to-face on Mondays and Tuesdays and remotely on Thursdays and Fridays and the other group does the opposite. Both groups meet remotely on Wednesdays and each class meeting lasts 51 min in either setting. My co-observer and I viewed Henry's face-to-face instruction with students via Google Meet remotely on consecutive Thursday-Friday sessions. Though 6 students meet face-to-face on those days, only 5 attended in person on Thursday and 4 attended in person on Friday. The IQA and TAG classifications for Henry's task are presented in Table 12 to provide context for the following sections.

Table 12

IQA Scores and TAG Classifications for Henry's Task

Score Level	IQA Score	TAG Classification	Mismatch
Phase 1: Selection	3	DM	No
Phase 2: Planning	3	DM	No
Phase 3: Setup	3	DM	No
Setup 1: CF	NA	NC	No
Setup 2: MR	NS	NC	No
Phase 4: Implementation	4	DM	No
Implementation: Student Work	3	NC	No

Analysis of Henry's Task. The task that Henry engaged his students with during the majority of the observed lessons was meant to be assigned as homework according to the MMR lesson plans. Instead, Henry implemented the task in class with his students. The task is one of many teacher resources provided by Desmos, titled "Point Collector: Lines" (Desmos, 2021). Like other "Point Collector" activities, the purpose is for students to "collect" coordinate points in the *xy*-plane by typing an appropriate inequality that contains such points in the solution set. For example, the point (0, 0) can by "collected" by typing the inequality $2x + y > -3$ because the origin is contained in the associated solution set. Throughout the activity, students progress from one slide to the next by typing or editing existing inequalities to collect as many highlighted coordinate points as possible. Collecting a blue-highlighted point grants a point, whereas collecting a red-highlighted point subtracts a point from a user's score; the goal is to obtain the optimal high score for each problem. The following paragraphs describe the portion of the Desmos activity that students engage in during the two observed lessons (some parts were assigned for homework afterward and not discussed during the two lessons).

The Phase 1: Selection IQA score for the Desmos task is a 3. Though the task does not explicitly prompt for the evidence of students' reasoning and understanding to warrant a score of 4, it engages students in problem-solving and requires the use of multiple representations. The task is completely open-ended and direct instructions for how to earn the maximum number of points for each problem are not provided. Students must use their knowledge of linear equations and inequalities flexibly, especially if they do not immediately identify an inequality that will earn the most points on their first attempt. Moreover, there may be infinitely many correct solutions for each problem, allowing students to earn the maximum number of points using a variety of solutions and solution strategies. Throughout their engagement with the task, students can explore how and why particular linear inequalities work to solve each problem (and why others do not).

Henry made no changes to the task when planning, resulting in an IQA score of 3 for Phase 2: Planning also. He used the same link to the task that is provided in-text, directing students to use the Desmos version of the task. Task setup, Phase 3: Setup, was also scored as 3 because Henry had his students begin working on the task without a preliminary discussion. In terms of the EIQA rubrics, the Setup 1: CF score is NA because the task does not involve a problem-solving (i.e., a real-world) component. Setup 2: MR was scored as NS because there was no whole class discussion of the task prior to student work on it.

Though the task itself received an IQA score of 3, task implementation scored a 4 because Henry elicited evidence of students' thinking and reasoning throughout their engagement. The general pattern for the lessons was that Henry (a) provided some individual student work time for a problem, (b) brought the class together to discuss that problem, and then (c) directed students to begin working on the next problem. Through his questioning during the whole-class conversations, Henry prompted students to make connections between algebraic and graphical representations for each inequality and to communicate these connections verbally. For example, when one student said that the statements $y > x$ and $x < y$ meant "the same thing," Henry echoed this idea back to the class and asked, “How can using the graph justify it?” Students responded that the associated graphs contained “the same shaded parts” and noticed that the inequality sign opened toward the same variable in either case.

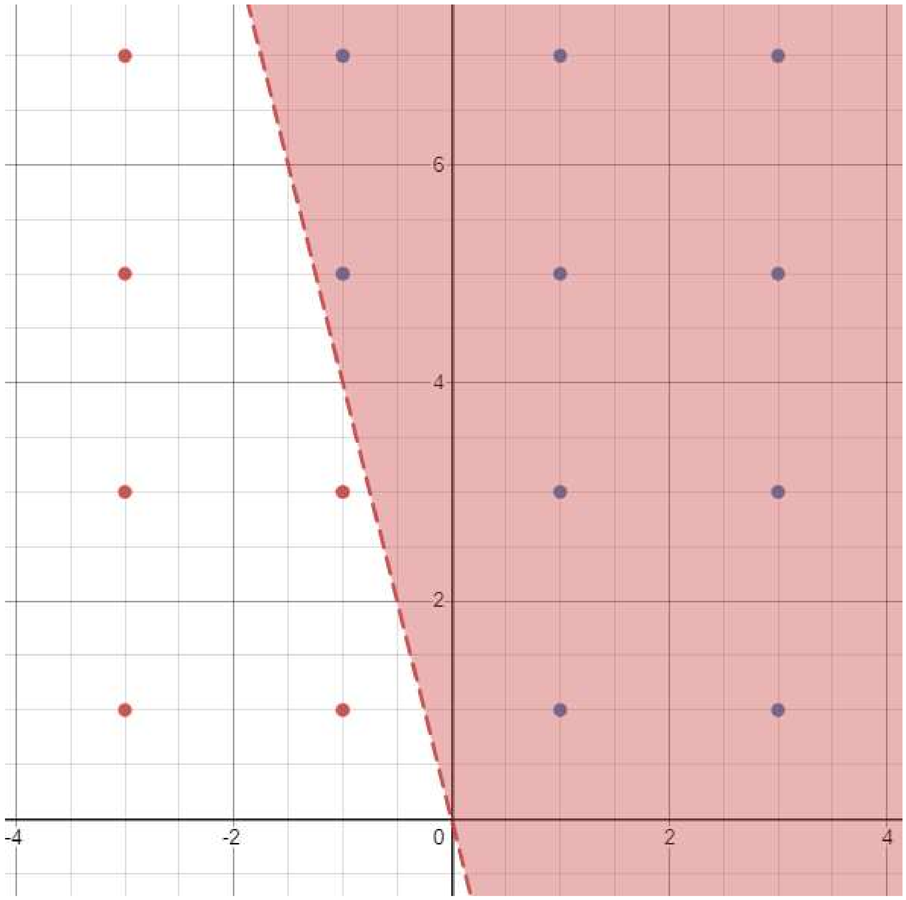
Henry additionally challenged students to determine whether their solutions were optimal, earning the highest possible score for each problem. After asking students what inequalities they used and projecting them on the smart board, he asked them questions such as, “Can you beat that?” and “Why can’t you do better?” These questions prompted students to explain why they could not obtain higher scores after collecting all the blue points and minimizing the number of red points. Henry developed an additional question to expand students’ thinking between the first and second lesson: he wrote the points (0, 2), (3, 2), and (0, 4) on the smart board and asked his students to identify a set of inequalities that would contain (shade over) all three points with the smallest area possible. Though students ran out of time to explore this question in great depth, they managed to conclude that the region with the smallest area would be the triangle with the above three points as vertices. Overall, teacher questioning and student explanations characterized the two lessons, resulting in an incline from an IQA score of 3 for task potential to a score of 4 for implementation.

The Implementation: Student Work score for Henry’s task is 3 for the same reasoning as in Phases 1–3. Though students were consistently prompted for explanations and verbally provided evidence of their thinking and reasoning, the online task itself did not provide students with opportunities to explain their work in writing. Henry downloaded and saved PDF copies of students’ responses to each problem, including only their final answers without written evidence of the thinking behind them. Each sample includes a graph of the points that students were to collect, the graphical representation of the inequality that a student typed as their final answer, and the solution set shaded on the graph. The student work also includes the algebraic representation for the inequalities that each student typed as their final answers (e.g., $y > -4x$). I replicated the graphical output for one student's work using Desmos, which is depicted in Figure 5.

Henry's Analysis of the Task. Because there were no changes for Phases 2–3 of the Desmos task, this section includes only Henry's analyses for Phases 1 and 4. His TAG classifications for the task at Phases 2 and 3 were identical to that in Phase 1: Selection and did not merit considerable discussion. Henry considered that the potential of the task (Phases 1–3) was at the level of *doing mathematics* primarily because the task was "open-ended," and students would be "exploring" throughout it. He explained that students would be "looking at trying to collect as many points as possible, and they have to do that mathematics by determining inequalities that will give them their best region." Students could not necessarily look at a problem and identify a correct solution to obtain the maximum number of points immediately; they would need to "try some things" to develop and test their conjectures. Students could also make mathematical connections, according to Henry, as they wrote algebraic inequalities and related them to the graphs that were generated by the Desmos graphing software.

Figure 5

Exemplar of Student Work for Henry's Graphing Task



Note. This representative sample of student work is a graph of the inequality $y > -4x$ in the *xy*-plane. The points (1, 1), (1, 3), (3, 1), …, shown in blue, are ones that the students were expected to "collect" to earn points. The points (–1, 1), (–1, 3), (–3, 1), …, shown in red, earned negative points and were not to be "collected." This is a recreation of one of Henry's students' work, earning a maximum possible score for this problem.

Implementation remained at the level of *doing mathematics* for Henry because students did "a pretty good job explaining things." He was satisfied with how they worked to solve the problems on the Desmos activity but also with their verbal explanations. For example, when one student suggested to use the inequality $y > -4x$ and another suggested to use $y > -\frac{2}{.5}x$ to solve one of the problems, Henry asked whose was "better?" In response, one student explained that they were both the "same" inequality, comparing the slopes –4/1 and –2/.5 and realizing that the first was just "twice as much work to do" as the other (i.e., the numerator and denominator of the first are twice the numerator and denominator of the second). I noticed how Henry praised his students' use of mathematical terminology throughout the two lessons when they used terminology such as "intersected" to describe how two linear inequalities crossed graphically. Henry was also pleased with the algebraic inequalities that students had come up with:

> They [the students] were really doing a good job of getting a lot of the blue dots… Three of them, especially, they were typing in an inequality, and usually their first inequality… accomplished pretty much what they needed to do as far as getting as many of the blue dots, getting the best score as possible… and it wasn't just a guess and check thing.

Henry acknowledged that his students were able to use their knowledge of lines to determine initial inequalities that would collect many of the blue points. Then, they simply adjusted their algebraic formulas as necessary to optimize their scores.

Phase 1: Task Selection. One of Henry's criterion for task selection is that the mathematics content "covers something that applies to the class," meaning that it addresses content that is relevant to what students are expected to learn. However, tasks must also encourage "some higher order thinking." Henry wants students to "know some material," but "it's more important that when they leave [my class], they're able to think and process problems through problem solving." He would rather spend more time engaging students in tasks that encourage mathematical sensemaking because if he can teach them how to think, then they can reason through problems they might not have seen before. Henry explained,

> If I can get them to think and work through problems, even if there's a few things I don't get to cover, at the end of the year, they'll be able to process through them and be able to figure out things on their own.

Engaging his students in tasks that focus on mathematical sensemaking is critical to his role as a mathematics teacher. To Henry, the "Point Collector: Lines" (Desmos, 2021) task is beneficial for students' learning because it allows them to expand their knowledge in a non-algorithmic way. Rather than simply memorizing an algorithm for graphing linear equations and inequalities, his students were challenged to make connections between algebraic and graphical representations in Desmos which would "help them remember things better."

Henry prefers active student engagement rather than demonstrating and lecturing to his students. "The less talking I do, the better," he explained, a philosophy that applies to MMR and his other classes. More so in MMR, but also in his other classes, Henry has his students engage in activities that are "a little non-routine," avoiding the monotonous classroom procedure of taking notes and working on practice problems every day. For example, he engaged his MMR students in a Desmos activity to introduce linear inequalities on the remote learning day that week. The task was open-ended and encouraged students to explore the Desmos graphing calculator as they graphed various linear inequalities and answered questions such as the following: “Why is it [the software] shading half of the plane?” and “What is the meaning behind it?” Moreover, Henry did not provide any clues or suggestions as to how to answer each question. “I want you [the students] to think about that,” he asserted. Similarly, he chose to engage his students in the “Point Collector: Lines” (Desmos, 2021) task so that they could explore and make mathematical connections on their own.

Henry admitted that he is not as “task-oriented” when teaching classes other than MMR. “Part of that is, well, a lot of that’s on me,” he confessed, explaining that he might feel even more “pressed for time” addressing the required course content if he did so. Henry thought that if he was “willing to sit down and do a whole lot more thinking,” he might be able to adopt the use of more mathematical tasks in these courses, especially in Precalculus and Calculus. The College Algebra course he taught through a local community college was “a little more dry” for the same reason: he is pressured to progress through the content of the course and does not have the flexibility he does with MMR. However, Henry aims to maintain the pedagogical approach he uses for MMR in his other courses. “For me, a good class is, I ask a question and just let the students have at it. And then they can ask questions, but they know 90% of the time I answer their question with another question.” Though his other courses incorporate fewer tasks, Henry avoids presenting his students with algorithms for them to regurgitate using different numbers. This is something that he has done throughout his entire teaching career.

Phase 2: Task Planning. When planning his lessons, Henry attends specifically to questioning. He anticipates two types of questions prior to instruction: (a) questions that he should ask "to get the students thinking" and (b) questions that his students might ask and how he can respond. For instance, he expected that his students might have difficulty conceptualizing and graphing linear equations because "they haven't done that for a while." Henry considered graphing an equation such as $x = 5$, explaining how he might ask a struggling student to plot various points where the value of x is 5. He would then ask, "Do you see a connection?" hoping that the student would realize that the set of points forms a vertical line. Henry thought that he might have preemptively addressed some of these issues by having his students complete the warm-up exercise on the remote learning day that week (the other Desmos task described in the previous section).

In anticipating how he might respond to students' questions, Henry planned to help students make connections to what they had learned before. He sequenced the two MMR tasks (the warm-up Desmos activity and the "Point Collector: Lines" [Desmos, 2021] task) in a way that his students might be able to make connections from one to the next. Hoping that his students had encountered a similar situation on the remote learning day, he found it likely that he would ask, "Does this have any connection with what you did yesterday?" However, he recognized a potential flaw with his plan, explaining that "Some of these kids are hit or miss whether they do the work on the remote day that day…some of them always manage to forget." Henry thought that it might also be helpful to refer his students back to what they had done in Algebra 1 and Algebra 2 if they struggled with lines and linear inequalities. He also hoped that his students might recall their work on previous tasks in the MMR course, such as “Follow the Bouncing Ball” and “Bungee Barbie,” which focused on linear functions.

Henry adapted his plans for the task in several ways during the observed lessons. He initially assigned the “Point Collector: Lines” (Desmos, 2021) task as homework, aligning with the MMR lesson plan for the Context. However, Henry realized that most of his students had not attempted it yet and decided that it would be beneficial for them to work on it in instead. He had also planned for his students to work on the task in small groups, but with only four students in class learning face-to-face each day, he decided to implement the task as one large group. Henry made this change so that he could “get a better feel” for his students’ understanding and he could ensure that every student provided their input and thinking. “I had a feeling that would work better as far as getting everybody involved,” he said.

Another adaptation, as I described previously, was the inclusion of the problem that engaged students in determining the minimum area between the points (0, 2), (3, 2), and (0, 4); this was a problem that Henry thought of the night before the second lesson and was not provided in the MMR materials. Henry felt that these spontaneous adaptations delayed the pace of his instruction but were worth the investment. “If I have to go that extra day, I'm going to go that extra day… had I not gone back over some of that stuff, then what was the point of doing it anyhow?” It was worth spending more time and helping students develop deeper understanding of the mathematics. As a mathematics teacher, “you’ve just got to be flexible,” Henry concluded.

Phase 3: Task Setup. Recall that there was no setup phase for Henry's task, as he simply directed his students toward what they should start working on and did not launch the task with a whole-class discussion. Therefore, his second interview focused entirely on the implementation phase of instruction.

Phase 4: Task Implementation. During task implementation, Henry focuses on promoting productive struggle and setting the expectation that making mistakes is a part of the learning process. He is focused on "trying to get them [his students] to try things," even if they are not correct on their first attempt. In our second interview, Henry elaborated on his interactions with one of his students who was hesitant to work on the Desmos activity. Initially, the student stared blankly at his laptop screen without attempting to enter an inequality for a certain problem. "I think he was pretty anxious he would do something and do it wrong," Henry explained, leading him to ask, "What are you going to do? What's going to happen if you put in something that doesn't work?" By doing so, he encouraged the student that he could always try something else and attempt to improve his thinking each time. Henry reflected that many of his MMR students likely felt the same way because they had not been successful with mathematics in the past. He said that "a lot of these [MMR] students end up liking math a whole lot better than they did before" because the course provides a space where they can feel "comfortable" making mistakes.

Henry questioned his students frequently throughout the implementation of the task. For example, he asked questions to elicit students' reasoning, such as, "Why did you put *y* first?" and "Why do you think so?" Henry asks numerous questions in his teaching for two key reasons: (a) as an assessment of what a particular student knows and their level of understanding, and (b) so that other students learn from each other's ideas. Addressing the first reason, Henry asked his students how they could use the graph provided by Desmos to justify that the statements $y > x$ and $x < y$ were equivalent; he said that he thought of and asked this question because he wanted to gauge his students' understanding of what they were doing and what each of the inequalities represents. He felt that many students, not just those in his MMR class, failed to understand that the graphical representation of an equation or inequality represents its associated solution set. Regarding the second reason, Henry would rather have his students explain ideas to each other than simply memorize what he tells them to do. He acknowledged that his students could explain the same things that he could, but other students might learn more from each other because their explanations may be in "terms that probably make more sense to them." Henry believes that students learn and retain more information when it comes from another student.

Asking frequent questions prompted Henry's students to provide evidence of their mathematical thinking and reasoning, ultimately raising the cognitive demand of the task from a 3 to a 4 on the IQA rubrics. Henry reflected that, "Having them explain makes them think more… and then it makes them have some deeper thinking about what they're doing," realizing that students think deeper and more analytically about their work when explaining it to others. He also acknowledged that students do not fully understand mathematics conceptually unless they can explain it logically and clearly, recalling the familiar student phrase, "I know what to do, I just can't tell you what to do." To Henry, students are not “cognitively… there yet” if they are unable to describe how they solve problems and cannot explain how they know they are correct. Henry held his students accountable for providing such evidence of their thinking and reasoning as they engaged with the Desmos task; he questioned one student, who “was waiting to see what somebody else said” on a particular problem, pressing him to explain his thinking and develop his own solution. “That wasn’t the whole point of the activity… to see what somebody did and then mimic it,” Henry point out afterward.

Isabel

Isabel has a bachelor’s degree in mathematics with a teaching license and has taken graduate-level coursework in mathematics as well. She has been with her current school for 3 years and has taught for 6 years in total. Isabel is the only one of the 9 participating teachers who does not teach MMR. However, she attended the Modspar ATC program at Ohio University more than once and therefore met the selection criteria described in Chapter 3. Her teaching position is unique because her students learn at their own pace; Isabel’s students work individually or in groups, when possible, as they access and complete assignments on Google Classroom. They progress from one lesson to the next upon completion, working through a checklist of tasks and assignments until they reach a test or quiz, after which they move on to the next topic. Isabel’s role as the teacher is to facilitate students’ self-engagement, direct and support them as they move from one subject to the next, and provide feedback and assessments as they advance.

Though her teaching environment differs significantly from other teachers participating in this study, I chose to include Isabel because of the qualitative interview data she could provide to answer the research questions. Through dialogue with Isabel, we identified one of her Algebra 1 classes to observe because some of her students were ready to begin a task that she selected and adapted from the internet. As I describe in the following paragraphs, observing her class was not ideal because not all students work on the same task at the same time. Moreover, there were limited interactions between Isabel and her students because of their unique arrangement. However, interviews with Isabel contributed data to answer the research questions of this study and observations provided additional interview questions that I had not planned in advance. My co-observer and I joined her Algebra 1 class remotely via Zoom. The following paragraphs describe the task that Isabel modified for two of her Algebra 1 students to complete during the observed lessons. I have also provided the IQA and TAG classifications for Isabel's task in Table 13 to provide a synopsis for the following sections.

Analysis of Isabel's Task. Isabel's task is unique because it did not originate from the MMR course materials. Instead, it is a task that she found online and modified for use in her own classroom. The task, titled "Fitting a Line to Data - Earnings and Educational Attainment" (United States Census Bureau, 2021) engages students in analyzing data pertaining to an individual's educational attainment (years of education and degrees earned) and their median annual income. Throughout the task, students are expected to identify patterns and describe data, create a scatterplot, create a line of best fit by inspection, and find the equation of the line and use it to answer questions about the real-world situation. The Phase 1: Selection IQA score for the task is 4 for several reasons: (a) the task involves making explicit connections between numerical, graphical, and algebraic representations as students analyze data, graph them, and determine an appropriate line of best fit; (b) the task requires students to identify patterns within the data and to develop generalizations based on those patterns; and (c) the task explicitly prompts for evidence of students' reasoning and understanding in various ways, such as: "Turn to your classmates and compare your lines and equations. Do some of your classmates' lines better fit the data? Explain" (United States Census Bureau, 2021, p. 11).

Table 13

IQA Scores and TAG Classifications for Isabel's Task

Score Level	IQA Score	TAG Classification	Mismatch
Phase 1: Selection	4	PwC	No
Phase 2: Planning	4	DM	No
Phase 3: Setup	NR	DM	No
Setup 1: CF	NR	NC	No
Setup 2: MR	NR	NC	No
Phase 4: Implementation	NR	DM	No
Implementation: Student Work	NR	NC	No

The Phase 2: Planning version of the task included formatting changes and adjustments to the work that students were required to do. One influential modification to the task was having students use Desmos to create a scatterplot and linear regression model rather than plotting data and drawing a best fit line by hand. Isabel also added questions focusing on independent and dependent variables and on correlation. She removed the prompt, “Turn to your classmates and compare your lines and equations. Do some of your classmates’ lines to better fit the data? Explain” (United States Census Bureau, 2021, p. 11),

because her students worked independently on the task. She replaced it with the following:

> Based on the equations, someone with a Bachelor’s degree makes more money than someone who only graduated high school. Does that mean that everyone with a Bachelor’s degree will make more than everyone with only a high school degree? Explain using complete sentences.

Isabel enhanced the requirements for students to explain their thinking in writing by adding phrases such as, “Explain using complete sentences,” “Show your work,” and “How do you know?” In summary, the Phase 2: Planning IQA score remained a 4 because the mathematics content of the task remained similar, and students were still expected to provide evidence of their thinking and reasoning in writing; perhaps even more so.

Phase 3: Setup, task setup, was not scored using the IQA and assigned “NR” due to Isabel’s unique teaching situation. Only two students worked on the task and did so at their own pace during the observed lessons; this setup was unique because not all the students were included at once. Therefore, assigning an IQA score was not reasonable because the setup of the task differed significantly from the others. Additionally, the setup of the task for each student occurred only briefly as Isabel provided each student with the handouts, explained her expectations for what they would submit to her, and reminded them to check in with her regularly so she could monitor their progress. These explanations lasted for only several minutes prior to each student's individual engagement with the task.

Similarly, the observed implementation of the task was not scored using the IQA and was coded as "NR." Because only two students engaged with the task during the lessons rather than the whole class, assigning a score similar to other teachers' tasks was inappropriate. Throughout the observed lessons, Isabel monitored her students' progress as they individually completed instructional tasks and assessments at their own pace. Most of her interactions with students consisted of asking questions such as, "Solving for *x*, what did you do?" and providing direct support (e.g., explaining what the term "outlier" means and providing an example of a cluster of data).

In one instance, Isabel helped a student to interpret the *y*-intercept of her regression model by asking a series of questions. In this instance, the *y*-intercept indicated that a person with 0 years of education would earn a nonzero income; nonsensical, but the purpose of the problem was for students to recognize this. Initially, the student misinterpreted the context and assumed that high school students had 0 years of education. In response, Isabel said "You have approximately 8 years of education. Who has 0?" to get the student to rethink her response. After concluding that a small child has 0 years of education, Isabel followed up by asking the next question on the handout verbally: "Is it useful to know what their income would be?" The student responded, “No,” prompting Isabel to confirm that the student was correct. Then, Isabel asked her to explain her thinking in writing. This exchange in dialogue was typical of Isabel’s interactions with students on a variety of tasks.

Because I only received one sample of student work from Isabel, I cannot reliably assign an IQA score for Implementation: Student Work. The sample indicates a moderate level of understanding on behalf of the student; they correctly identified the independent variable as “education” and the dependent variable as the “income,” but could have been more specific about how “education” and “income” were defined. For example, “income” could have been written as “median household income, annually.” Moreover, the student noticed the “strong positive trend” in the data, indicating that an increase in education is associated with an increase in income, but did not explain this connection in detail. Their interpretation of the slope lacked substance, but the student made the connection that a negative *y*-intercept means “You lose money if you have no education.” When prompted to write about what the student learned, she commented, “I learned that Desmos does all the work,” among other general statements. The student’s numerical calculations are correct, but it seems that they might be lacking a complete understanding of the concepts underlying the technology she used and how the real-world situation related to the mathematics.

Isabel’s Analysis of the Task. Isabel classified the original version of the task as *procedures with connections*. She referred to some of the procedures that students were expected to do, such as “plotting their coordinate points and when they’re using their ruler to then draw their line, that’s all procedure stuff.” Isabel also noted that the task suggests pathways for students to follow from one problem to the next and requires some degree of cognitive effort: students "have to think back to, 'how do I find slope,' especially because that's two points on their line, not necessarily two data points." However, the task was not at the level of *doing mathematics* according to Isabel because students are "connecting what they would have learned for line of best fit… but not necessarily doing the analyzing" that would be required at the highest TAG level.

Initially, Isabel stated that her modified version of the task remained at the *procedures with connections* level but contained more elements of *doing mathematics* than the original. She thought that her version of the task required more analysis and interpretation than the original and required "much more self-monitoring and self-regulating." Isabel later explained that her task was "close to 50-50" between *procedures with connections* and *doing mathematics* because it provides a specific pathway for students to follow (e.g., running linear regression and using a formula) but also prompts them to interpret slope and the *y*-intercept in terms of the real-world context. The more that Isabel explained, the more she felt that the task was at the level of *doing mathematics*; this was, in her mind, because students would have to access relevant, prior knowledge about linear functions and apply it to a unique situation.

Because the setup of the task was minimal, Isabel also noted that the cognitive demand of the task did not change from *doing mathematics*. She said, "I guess I do point out that this ties in regression, but I don't necessarily say, go back and look at your last set of notes and look at all of that stuff." In her typical instructions to students, she only explains how this task is "more of an activity," different from other lessons that students complete online. She instructs students to check in with her periodically so that she can determine whether they are on the right track but does not think that this lowers the cognitive demand of the task.

Isabel attributed the *doing mathematics* nature of task implementation to the work and written responses that students provided on their task handouts. One of the students I observed completing the task had an IEP and required “a couple of things accommodated for” because of the amount of reading and writing involved in the task, Isabel explained. However, she also said, “I don’t think she had any issues. She had me check over stuff and I don’t recall anything being wrong. Or maybe one thing wrong as opposed to maybe a couple of things in the first section.” At one point during the lessons, the second student working on the task struggled using her linear regression to predict the income for someone with only a high school diploma. Instead of assisting the student directly, Isabel had her talk to the other student who had already completed that part of the task, and the two students discussed the problem together. Isabel felt that these interactions helped to maintain the cognitive demand of the task because her students succeeded in solving the problem without her direct intervention.

Phase 1: Task Selection. Isabel’s selection of tasks is somewhat limited because of her unique teaching situation. As I described earlier, her school is formatted so that students learn at their own pace individually by progressing through materials in Google Classroom. Isabel has some flexibility in how she uses and adapts these materials, if she follows her school’s curriculum map: “It doesn't tell us when we're supposed to teach stuff and it doesn't tell us what exactly to teach,” she explained, but many of the materials were provided to her when she started teaching at her school and she continues to use them. The Google Classroom materials for Algebra 1 appear to be typical of the traditional Algebra 1 course, focusing on purely abstract, mathematical concepts. For example, the resources that she shared and discussed with me during our first interview addressed the equations and graphs of lines.

The “Fitting a Line to Data - Earnings and Educational Attainment” (United States Census Bureau, 2021) task was one of several “enrichment activities” that Isabel had her students complete. She tries to use such an activity each unit, though she was unable to use some of them because of safety precautions involving the COVID-19 pandemic. Isabel used a different real-world activity involving a restaurant’s production of guacamole to strengthen students’ understanding of linear regression in the previous year; it was “still a good connecting lesson,” Isabel thought, but “the students… didn’t really care” about the real-world aspect of the task; this year, she sought an activity addressing similar mathematics content with a more relatable, interesting real-world connection. She felt that the “Fitting a Line to Data - Earnings and Educational Attainment” (United States Census Bureau, 2021) task was more interesting for her students because they all receive an education “to some extent,” though some might choose to pursue higher education and others might not. “What's better about this one,” Isabel reflected, “is that it's actually something that they [students] care about. It’s relevant.” After witnessing several students complete the task, she confirmed that she favored the change from last year’s task and would continue to use the newer task in the future.

Because of her unique teaching situation, I also inquired about some of the tasks that Isabel had used when teaching under more typical circumstances at a previous school. She explained that the Modspar PD at Ohio University had been influential in her use of mathematical modeling tasks when she had the ability to engage all her students at once. Such tasks involved the open-ended exploration of real-world situations: previously, Isabel had her Geometry students explore the amount of pizza they could buy from various companies if each of them brought in a fixed amount of money. She felt that the use of modeling activities from Modspar was "beneficial" for her students because "it was something that they seemed to enjoy." Moreover, her students were able to experience "a different way of learning and a different way of showing what they've learned." Isabel also appreciated how the open-ended nature of modeling tasks allowed her students to explore in various directions and helped them to gain confidence in developing and sharing their own ideas. If she changed back to a teaching format that enabled her to use more of these tasks in the future, she "would definitely" do so as she was able to.

Phase 2: Task Planning. As I described previously, Isabel made several changes to the "Fitting a Line to Data - Earnings and Educational Attainment" (United States Census Bureau, 2021) task. She chose to have her students perform linear regression using Desmos rather than sketching best-fit lines by hand. Isabel made this adjustment because she thought it would be difficult for her students to draw the line without her direct supervision to be sure that it was indeed "the best" line possible; having a less-optimal line would influence students' engagement with the remainder of the task, so

Isabel decided to incorporate Desmos instead. However, she emphasized that her students would still be responsible for the mathematical interpretation of their regression lines, such as determining the slope, *y*-intercept, and the correlation coefficient. Isabel included follow-up questions for her students to interpret this information in terms of the real-world context because it was absent from what she found online: "It [the task] asked about… what the *y*-intercept is and what it means, but it didn't really ask about the slope or the correlation coefficient," she explained. She also thought that it was important for her students to be able to use their regression equations to predict an annual income using a given amount of educational attainment; therefore, she composed additional questions to address this concept.

Phase 3: Task Setup. The setup of Isabel's task was unlike the 8 MMR teachers' because she provided individual instructions to each student as they prepared to work on it. Instead of a whole-class discussion, Isabel briefly explained her expectations for students' engagement with the task. This was their first activity for the year, so they would be "unfamiliar" with the setup. As she gave each student the task handout, Isabel explained that she wanted all their answers to be correct before submitting it. "Because of that, they should make sure to check in with me pretty regularly, have me check their work," she elaborated. She indicated which questions were "right or wrong type of questions," those with a specific, correct answer, and which were "opinion questions," those that involved written sentences and could be interpreted in various ways. Isabel instructed her students to check in with her throughout their work on the task because many questions repeated, except with different data, and some questions were based on students' responses to those prior; in both cases, she wanted her students to stay on the right track and not have to redo much of their work based on a single incorrect response.

Phase 4: Task Implementation. Isabel's primary role when implementing a task is to facilitate students' engagement. The nature of her interactions with students varies from day to day, but typically involves walking around the classroom and observing how her students progress from one assignment to the next. Most instructional material is provided in Google Classroom in the form of videos, created by Isabel or the other Algebra 1 teacher, though she sometimes provides one-on-one instruction when it cannot be done through video (e.g., when she plans to ask specific questions and wants to hear her students' responses). When students watch instructional videos and work on practice problems, Isabel spends more time observing; however, she becomes more active when her students ask questions about the problems they are working on. Isabel assesses students through short quizzes when they reach specific milestones, which she grades in class so that she can provide instant feedback.

One of the most frequent instructional moves that Isabel made during the observed lessons was providing support to struggling students. Some support was more direct, as she would either walk a student through a problem step-by-step or offer mathematical information that would help a student progress further. For example, Isabel told one student (not working on the task described in this section) that the procedure to find x-intercepts was to "set y equal to 0 and then solve for x." She provided a similar algorithm for identifying a y-intercept in response to the student's question. Other support was less straightforward: when the second student working on the "Fitting a Line to Data

- Earnings and Educational Attainment" (United States Census Bureau, 2021) task asked a question, Isabel instructed her to talk to the first student, who had spent more time working on the task. The amount of support she offers varies based on (a) the student she is working with and (b) the lesson that the student is working on. Isabel tries to avoid "giving them [her students] too much," but sometimes finds it necessary if a student is unable to find success after less-direct intervention. "Based on their math ability," Isabel is likely to respond differently to two students who ask the same question. Some lessons that introduce new topics "need that little bit more guidance" compared to others that might connect to ideas that students have already learned.

Isabel also asked numerous questions to her students throughout the observed lessons, such as "Solving for *x*, what do you do?" and "Where did you get 6 and 6 from?" Such questions gauge her students' understanding and help Isabel determine what they are thinking in a specific instance. One question that she asked a student working on the "Fitting a Line to Data - Earnings and Educational Attainment" (United States Census Bureau, 2021) task was for the student to interpret the meaning of the linear regression model she found using technology. Isabel asked, "Does that mean that a person with a bachelor's degree will always make more money than someone with only a high school diploma?" The student's initial response was, "Yes," but Isabel questioned her further to help her realize that this was not always the case. In general, Isabel asks follow-up questions when appropriate because they "make them [her students] think," whereas direct responses simply provide answers and encourage students to move on without thinking further. “I try to give them [students] more of those leading-type questions when possible, to help keep it more on them thinking than on me thinking,” she explained.

Cross-Case Analysis

In the following sections, I describe the 18 emerging themes that arose from the analysis of teachers’ interviews. The themes are organized into four groups as follows: (a) 5 pertain to the selecting of tasks from source materials, (b) 5 pertain to planning around the use of tasks, (c) 3 pertain to setting up tasks for student engagement, and (d) 5 pertain to implementing tasks with students. The following themes are based on teachers’ responses about the general use of instructional tasks and the use of the specific tasks that were observed as a part of the present study.

The Emergent Themes for Task Selection

The emergent themes for task selection are presented and described in Table 14. These themes answer the first part of research question 1, referring to teachers’ task selection and reasons for doing so: (a) teachers consider face-to-face, remote, and hybrid learning formats; (b) teachers promote active student engagement; (c) teachers address important mathematics content, skills, and processes; (d) teachers make connections to real-world contexts; and (e) teachers consider past success. The following paragraphs include a synopsis of how various teachers exemplified each theme.

Table 14

The Emergent Themes for Task Selection

Theme	Description
Teachers consider face-to-face, remote, and hybrid learning formats.	• Selected tasks are suitable for remote and hybrid instruction. • Selected tasks can be implemented safely in face-to-face settings.
Teachers promote active student engagement.	• Selected tasks are open-ended and encourage exploration. • Selected tasks involve group work and collaboration.
Teachers address important mathematics content, skills, and processes.	• Selected tasks address content that teachers find important for students to learn. • Selected tasks emphasize problem-solving and critical thinking.
Teachers make connections to real-world contexts.	• Selected tasks include authentic real-world scenarios. • Selected tasks exceed the level of traditional word problems.
Teachers consider past success.	• Selected tasks have been successful in previous years. • Selected tasks have been successful with other teachers.

Teachers Consider Face-to-Face, Remote, and Hybrid Learning Formats. As shown in Table 15, Adam, Beth, Debbie, and Ethan explained that their hybrid and remote teaching environments influenced their selection of tasks. When selecting a task, Adam wonders whether his students will be engaged and make connections. He acknowledged that his students might access instructional tasks differently based on whether they are learning remotely or face-to-face. “I have kids… in class [MMR] that join virtually and those that are in my classroom at the same time… can I make this fit for both groups?” Similarly, Debbie reflected that some tasks might not be suitable for students who attend school virtually or for those in hybrid environments where students are split into face-to-face and remote groups. She explained that teaching the MMR course in a fully remote setting would be more difficult than teaching it in a hybrid or entirely face-to-face setting.

Table 15

Teachers Consider Face-to-Face, Remote, and Hybrid Learning Formats

Teacher	Supporting Interview Quotation
Adam	Is it [the task] something that can meet the hybrid standard that we're at right now? Because I have kids in this class age that join virtually and those that are in my classroom at the same time. So, time, and can I make this fit for both groups?
Beth	There's only a few lessons from each theme that I can do strictly online easily. So that [the task] was definitely one of the ones that I could do online easily.
Debbie	Having this MMR class to be here hybrid right now has helped to facilitate some of these activities where, if we were fully virtual, it would be hard.
Ethan	I'm looking for things that might translate well to remote learning, especially because we've been remote… all but for 3 weeks.

Beth chose to use the "Discovering Slope" task because her students could engage with it online and did not need to be with her in the classroom face-to-face. Aside from one station involving a physical ramp, the remainder of the stations provided the necessary information for students to engage with them readily. Ethan specifically chose tasks that could "translate well," or be easily adapted to, remote learning after his district shifted from hybrid learning to fully remote instruction. Though his task required students to collect data by bouncing elastic balls, another MMR teacher had already recorded videos that students could watch to take measurements instead. Ethan considered himself fortunate to be provided with some materials that had already been adapted by other MMR teachers for remote and hybrid learning.

Teachers Promote Active Student Engagement. Table 16 shows that the 8 participants teaching the MMR course emphasized active student engagement as a motivator for selecting tasks to use with their students. Beth, Cathy, Henry, and Gwen specifically stated that they preferred engaging students using mathematical tasks as opposed to lecturing. For example, Beth desires "interactive and hands-on" lessons because she wants students to be "interested online rather than being lectured to." Similarly, Cathy, Debbie, Ethan, and Fred chose "hands-on" tasks to engage their students in seeing, feeling, and experiencing the mathematics for themselves. Debbie, who is also a science teacher, uses "a lot of hands-on stuff, especially at the beginning. So, it's exploration that they [students] can touch in some way… something where they can do some form of an experiment," as she does when teaching science. Ethan tries to incorporate more “hands-on” learning, group projects, and tasks in his other mathematics classes in addition to MMR.

Table 16

Teachers Promote Active Student Engagement

Teacher	Supporting Interview Quotation
Adam	We're project based where we're constantly doing activities and talking through and doing fun things.
Beth	We designed lessons that were very interactive and hands on. So, because we want them [students] to be interested online instead of just being lectured to.
Cathy	It's always been about engagement… If I can limit the amount time that I'm standing in the room yammering, putting kids to sleep, that would be lovely.
Debbie	I do a lot of hands-on stuff, especially at the beginning. So, it's exploration with something that they can touch in some way.
Ethan	I've tried to incorporate more hands-on learning and more group projects and more of this like, task idea, into my other classes.
Fred	I do a lot of investigation.
Gwen	If appropriate, tasks and activities should engage all students, not just the ones who “get it” already.
Henry	The less talking I do, the better, I always feel.
Isabel	I feel like it [modeling activities] makes it much more fun and enjoyable for the class period, as opposed to just sitting and learning and doing notes and doing a practice worksheet and taking a quiz and all of this.

The 8 MMR teachers chose tasks encouraging student exploration throughout the year. Beth, Cathy, and Gwen select tasks that intentionally "scaffold" students from one question or idea to the next. For these teachers, such tasks should be somewhat self-guided for students and allow them to progress independently or in groups by providing guidance but not "spelling it out," according to Beth. In Gwen's words, the task "tells them [students] what to do but doesn't tell them what to do." The task "tells them what to do" by offering instructions and directions but "doesn't tell them what to do" in terms of the specific mathematical processes and procedures they should use to solve a given problem or produce an expected result. This is evident in the tasks these three teachers chose for the purpose of this study: for example, Cathy's "Starbursts™ Grab" task guides students through an exploration relating hand size to the number of candies that can be grabbed by prompting students to consider relevant variables, to decide how they might be related, to collect data to provide evidence of this relationship, and to draw conclusions from the results. However, the task does not specify which variables to choose, how to measure them, or what type of model will best describe the potential relationship (though the placement of the task in the curricular materials suggests that a linear relationship is expected).

Several teachers mentioned that the ideal task is accessible in addition to being engaging. Gwen prefers to use tasks that she thinks engage *all* students, not just the ones who "get it," believing that such tasks should be "approachable" even for students with low mathematical self-efficacy and those with limited background knowledge. "For those students especially, tasks should make them feel comfortable enough to start." Gwen called such tasks "low-floor, high ceiling" tasks, meaning that they allow for "multiple entry points" and are accessible for students but also contain rich mathematics. Beth's description included tasks that allow students to brainstorm and "get their feet wet with the material" initially as a means of being accessible to many students. She selects tasks that ideally provide "some place that the kids could start," aligning with Gwen's notion of "low-floor, high ceiling" tasks.

Teachers Address Important Mathematics Content, Skills, and Processes. The 9 participating teachers highlighted aspects of their selected tasks that they felt were important for students to learn. Supporting evidence for this theme is displayed in Table 17. For some of them, the mathematics content was important. The three teachers who used the "Follow the Bouncing Ball" task, Debbie, Ethan, and Fred, desired for their students to understand the meaning of a functional relationship relating a dependent variable and its associated independent variable. Adam, Ethan, Henry, and Gwen explained that they might use each of their tasks in other courses, such as Algebra 2, because the mathematics content was relevant.

More so than the mathematics content, the teachers addressed mathematical skills and processes that are important for their students to learn through the tasks they chose. Cathy prioritized supporting students' ability to make predictions, whereas Gwen and Henry aimed to enhance students' ability to think critically about mathematical situations and solve problems. Debbie felt the strongest about the use of data to inform students' decision-making. Cathy and Fred stressed the importance of learning to use Excel spreadsheets, but Debbie and Ethan explained that the time to learn technology-specific skills may interfere with the mathematics. Instead, they made adaptations so that their students might still learn to work with data using more intuitive technology.

Table 17

Teachers Address Important Mathematics Content, Skills, and Processes

Teacher	Supporting Interview Quotation
Adam	It’s [the task is] dealing with converting measurements as well… I would use this in any of the really, Algebra 1-type classes.
Beth	It's important that they [the students] understand what slope really is.
Cathy	Another thing I always like to try to build into activities I make myself, that aren't always there in textbook activities, is predict first.
Debbie	I want to make sure that they [the students] are being challenged into thinking, “how can they support their reasoning, especially with data?” I think we overlook that in math a lot.
Ethan	It's a good task to get kids to kind of get the idea of this functional relationship.
Fred	I don't need them to do symbolic manipulation. The mathematics that I need them to do is to think about “What does this mean?”
Gwen	It’s [the task is] realistic. They [students] may not need to calculate slope in the future, but they will need to be able to approach real-world problems and think critically through them.
Henry	I want them [the students] to know some material. But I think it's more important that when they leave, they're able to think and process problems through problem solving.
Isabel	In this one [the education and income task] they [the students] actually had to run a regression… the last one [the avocado task] had just been like approximate using their line of best fit.

Teachers Make Connections to Real-World Contexts. Of the 9 participating teachers, 8 chose tasks that allowed students to connect mathematics to authentic, real-world contexts. The only teacher whose task had no real-world component was Henry, though he discussed the topic during our interviews. Conversely, Fred did not voice his thoughts about real-world applications despite choosing such a task, as shown in Table 18. Adam, Beth, Ethan, Gwen, and Isabel explained how the tasks they selected were relevant to students' lives. By having her students visualize slope in different ways, Beth hoped that her students would recognize and think about it when they encounter stairs, hills, and other real-life objects in the world around them. "A lot of these students that are in my class haven't seen how math affects their daily lives," Ethan speculated, hoping that his students might realize the relevance of mathematics through various MMR tasks.

Cathy, Gwen, Henry, and Isabel suggested that using tasks with meaningful real-world applications enhances students' engagement and involvement. Isabel replaced a linear regression activity with the "Fitting a Line to Data - Earnings and Educational Attainment" (United States Census Bureau, 2021) task because the former involved restaurants and avocados, far less pertinent to students than their future education, careers, and financial statuses. Adam recognized that many of his students may someday be involved in buying or renovating houses and that some of the real-world knowledge that they discovered through the task might be beneficial for them later in life.

Table 18

Teachers Make Connections to Real-World Contexts

Teacher	Supporting Interview Quotation
Adam	I have not yet heard the question, and I get it frequently, “When am I going to ever use this [math] again?” It has been taken completely out of the equation [in MMR].
Beth	One reason that our kids struggle so much, especially in Algebra 2, with math, is they don't see anything outside the math classroom. And I'm not sure a lot of math teachers do, either.
Cathy	It's not just mindless numbers, it's actual stuff… so it keeps their [students’] interests.
Debbie	I've always done that [student exploration] from year 1. But I think even more so now, where I’m tying it more to a task for a real situation beyond just exploring patterns.
Ethan	I feel like a lot of these students that are in my [MMR] class haven't seen how math affects their daily lives… They just don't see the importance of it whatsoever. And I think when they can see that connection, it encourages them to do that much more, and to do that much better.
Gwen	The ramps task connects math content to a real-world problem.
Henry	The more we can do where it’s not just something abstract. Where they're [students are] actually touching something or seeing something real-world or a little more real-world, the more they're into it, the more they're involved, and the more they think.
Isabel	What's better about this one [task] is that it's actually something that they [students] care about. It’s relevant.

Teachers Consider Past Success. As shown in Table 19, 6 of the teachers consider the past success that either they have experienced with a task, or, in Adam's case, the success that other MMR teachers report in their regular meetings. Beth, Debbie, and Ethan expressed that each of their tasks were successful in the previous school year for various reasons. Though Debbie was absent for a surgery and the "Follow the Bouncing Ball" task was implemented by a substitute teacher in 2019–2020, she learned afterward that her students "loved" the task and enjoyed it. Both Debbie and Ethan chose to reuse the task for the 2020–2021 academic year because the data collection procedures were simple and the linear relationship among the data was clear. According to Debbie, such data helps students to develop knowledge about linear relationships that might not be as evident with "roughly linear" data where the relationship appears to be "forced." Beth and Fred have used each of their respective tasks more than once in the past with success, making the use of each task a natural choice, especially considering the additional challenges imposed by the COVID-19 pandemic. The use of familiar tasks allows teachers to concentrate on other aspects of their instruction aside from content and pedagogy.

Unlike the other teachers, Adam expressed careful consideration of tasks that have been successful with other MMR teachers during the school year. With each teacher at a different pace, it should be expected that some teachers progressed more quickly than others through the MMR course. This was likely amplified by the fact that the MMR teachers were allowed more flexibility when choosing which tasks they would use due to the restrictions of the COVID-19 pandemic, leading some to advance through the materials more quickly based on which tasks they kept and which they skipped. At each of their regularly scheduled meetings with ODE, Adam took note of tasks that had been successful with other teachers and which had not, influencing his decisions when selecting tasks for his own classroom. Adam may not be the only MMR teacher who does this; naturally, teachers may be more inclined to select tasks that have been successful with others than those that have not.

Table 19

Teachers Consider Past Success

Teacher	Supporting Interview Quotation
Adam	We talk about lessons that are successful, what lessons they [MMR teachers] avoided. And then I sit back, and I make assessments: "Okay, should I even try this?"
Beth	This is the third time I've done it ["Discovering Slope"] and it works beautifully.
Debbie	I love this one ["Follow the Bouncing Ball"]. Last year when I did it, they [the students] loved this one.
Ethan	I think back to what worked well last year and what didn't… I remember, this activity was a really good activity.
Fred	I have a lot of experience doing it ["Follow the Bouncing Ball". So, I facilitated it over the years and developed it, I think a lot, over the years.
Gwen	I had to modify it ["Ramp It Up"] quite a bit last year. But having experience with most of last year's tasks, I am now reflecting on last year's ahead of implementation with the current class.

Summary of Phase 1: Task Selection Findings. Within the task selection phase, Phase 1, the MMR teachers selected tasks that fit the design and goals of the course. The themes involving active student engagement, mathematics content, skills, and processes, and connections to real-world contexts are evidence of this: as a course in mathematical modeling, the inclusion of tasks with genuine and interesting real-world contexts is an integral part of the course. Similarly, the PD for MMR teachers focused on student-centered pedagogy, including the use of group work, collaboration, and problem-solving with open-ended tasks. Though the scope and sequence of the MMR course activities had been prescribed by the Ohio Department of Education (ODE), the MMR teachers were allowed more flexibility in choosing which tasks they would select for instructional use and which they would not because of the COVID-19 pandemic. Unlike a typical year, in which the teachers would be expected to implement all the tasks, these 8 MMR teachers selected tasks that they thought might work best in their specific environment (e.g., face-to-face, remote, or hybrid). The teachers also used their knowledge and prior experience with the course to determine which tasks might be effective in their COVID-19 environments and which might be more difficult to implement, given the unprecedented teaching environments they were put in. Those teachers who were new to teaching MMR considered the expertise of others who had already tested some of the activities in their own classrooms.

The Emergent Themes for Task Planning

The emergent themes for task planning are presented and described in Table 20. These themes answer the second part of research question 1, referring to teachers' task planning and reasons for doing so: (a) teachers are flexible; (b) teachers consider goals and objectives; (c) teachers are faithful to the provided MMR lesson plans; (d) teachers anticipate challenges and student responses; and (e) teachers adjust for face-to-face, remote, and hybrid learning formats. The following paragraphs include a synopsis of how various teachers exemplified each theme.

Table 20

The Emergent Themes for Task Planning

Theme	Description
Teachers are flexible.	• Teachers adapt their plans daily. • Teachers allow students to drive the length of tasks.
Teachers consider goals and objectives.	• Teachers modify tasks based on learning goals. • Teachers anticipate how they will help students meet their goals.
Teachers are faithful to the provided MMR lesson plans.	• Teachers follow MMR lesson plans closely. • Some teachers deviated from lesson plans more than others.
Teachers anticipate challenges and student responses.	• Teachers predicted difficulties with mathematics content. • Teachers considered how they might respond and question their students.
Teachers adjust for face-to-face, remote, and hybrid learning formats.	• Teachers converted handouts into collaborative slideshows. • Teachers provided safe data collection procedures in face-to-face settings.

Teachers Are Flexible. As shown in Table 21, 7 of the teachers expressed flexibility when planning their instruction. Adam, Cathy, Debbie, Ethan, and Gwen modify their plans daily, depending on how a task unfolds on a given day. Though Gwen plans her lessons for the "big picture," including an entire unit at a time, she adjusts her plans according to how her students progress each day. Ethan, on the other hand, has been "burned one too many times" with planning too far in advance, inspiring him to plan smaller sections at a time and adapt as necessary. Debbie provided a specific example: she considered using Part 2 of "Follow the Bouncing Ball" if her students "flew through" Part 1 and they had extra time. Similarly, Cathy considered having her students collect data on Day 1 of the "Starbursts™ Grab" task if they finished their preliminary discussions early.

Ethan and Henry commented specifically on flexibility with struggling students, each confident that slowing down and taking the time to support them was worth it. Ethan will not move on from one task to another if his students have not completed it; he would rather exceed the number of days he planned for the completion of a task so that his students are not forced to move on without sufficient understanding. Similarly, Henry will "go that extra day" because he sees no point in rushing through a task if his students develop only a surface-level understanding.

Table 21

Teachers Are Flexible

Teacher	Supporting Interview Quotation
Adam	We're at a point to where we've done it [remote teaching] so many weeks now… You just go with the flow.
Cathy	If they [the students] finish early, I can always have them start gathering data.
Debbie	If they [the students] just really fly through this, which I don't think they will, but if they do, then I might do some modified version of the second part.
Ethan	I'm a big fan of slowing down when kids are struggling, and so I have this planned out for what I think it's going to look like tomorrow, but if we don't get through it all, then I'm going to pick up where I left off on Monday.
Fred	I guess since I've done this activity a lot, I feel real comfortable varying it a little bit.
Gwen	I modify my plans on a daily basis, but always keep my end goals and outcomes in mind.
Henry	If I have to go that extra day, I'm going to go that extra day because, had I not gone back over some of that stuff, then what was the point of doing it anyhow?... You just got to be flexible.

Teachers Consider Goals and Objectives. Beth was the only teacher who did not mention a goal or objective that she either (a) considered generally when planning or (b) made a specific adjustment to a task to address, as shown in Table 22. Adam, Gwen, and Isabel communicated their general attention to standards and objectives when planning: Adam referred to the MMR lesson plans, which include the CCSSM Standards for Mathematical Practice (NGA & CCSSO, 2010) that are addressed for each Context. The lesson plans also include objectives established by ODE. Gwen's reference to "standards and objectives" was ambiguous and she did not specify which standards she referred to, though it is possible that she was also attentive to those provided in the MMR lesson plans. Isabel's use of the word "standards" is different, but also unclear; however, she may have been referring to Ohio's Learning Standards for Mathematics because these standards are relevant to high school Algebra 1.

Debbie, Ethan, Fred, Gwen, Henry, and Isabel modified each of their tasks to better address their own goals and objectives for their students. For Debbie and Ethan, this involved the use of technology, as each expressed that the focus on the mathematics may be lost if technology became the focus of the task. Debbie chose to use Data Classroom rather than Excel because it was more intuitive, whereas Ethan provided data tables for his students so that they could spend more time focused on the mathematics than on how they should present their data. Gwen removed the first page from the "Discovering Slope" handout (not the task I observed, but still important to mention) because she wanted her students to ask for information rather than providing it herself. Fred and Isabel added questions to each of their task handouts, prompting students to interpret the meaning of the slopes and *y*-intercepts in their regression models. Similarly, Henry created and posed an additional problem that was not originally part of his task to verify that his students possessed an understanding of the mathematics involved. Cathy, on the other hand, realized after the implementation of the "Starbursts™ Grab" task that she meant to revise the "yes or no" questions on the MMR handout.

Table 22

Teachers Consider Goals and Objectives

Teacher	Supporting Interview Quotation
Adam	I pull up… the lesson plan, and I see it's pretty detailed in terms of the goals and objectives that the lesson is intending to achieve.
Cathy	We're in a technological world. Are we wasting our time making them find the equation anyway? I mean, are they going to have a computer that can make the equation for them?
Debbie	There's a Part 2 that I think kind of deviates from the original goals of what are described there [in the MMR lesson plan]. And it seems to go off in another direction.
Ethan	I know that they [ODE] wanted to really stress the spreadsheets and like, the Excel stuff, and I'm like, "We could do it, but that's going to take another day or two of just talking technology as opposed to getting at the math behind it."
Fred	I do want everybody to sort of become well versed in the spreadsheet use… I really would like everyone to do a spreadsheet.
Gwen	I have a general understanding of what will be covered, what tasks and materials I will use throughout the unit, and what standards and objectives the students will be expected to master.
Henry	I was like, "Well, how can I extend this even beyond what they're doing and also make sure they're showing me that they really understand how to get the equations for the lines?" So, I came up with that problem.
Isabel	Describing or interpreting the slope, intercept, and the correlation coefficient, I just thought that was useful because that's something that is in our standards that we talk about multiple times… So that was something I felt was important that needed to be added that wasn't in the original [task].

Teachers Are Faithful to the Provided MMR Lesson Plans. Though some teachers communicated a greater dependence than others, 6 of the 8 MMR teachers referenced the lesson plans provided by ODE to some degree. A sample quote for each is listed in Table 23. Of the 6 teachers, the 2 teaching MMR for the first time (Adam and Cathy) conveyed the greatest reliance and fidelity to the provided lesson plans. Adam explained, "I have to go through that [the MMR lesson plan] and see what's required of me and what's something that I should leave in the hands of the students," using the MMR lesson plans to influence his role as the teacher and the role of his students. Cathy's statement, "I try to stay true to it [the lesson plan], whatever they [ODE] want the lesson to look like" indicates her desire to implement each lesson in the way that ODE expects it to be implemented. She also attributed the lack of a "good enough post-discussion" in the MMR lesson plan as one of the reasons that she did not conclude the task with a whole-class discussion, suggesting even an overreliance on the materials.

Beth, Debbie, Ethan, and Fred, teachers with more experience teaching the MMR course, made mention of the provided lesson plans but did not reveal the same level of reliance as Adam and Cathy. Alternatively, Debbie and Fred indicated some level of comfort in deviating from the prescribed MMR instructional procedures. Debbie felt that some of the prompts on the "Follow the Bouncing Ball" handout were unnecessary and that she could lead her students toward these ideas through discussion; for example, she explained that her students "don't need all those words there" defining and telling them to use rebound height as one of the variables they would measure. However, she recognized that such suggestions were meant to "help with the implementation, so the task does what it’s supposed to do.” Likewise, Fred’s years of experience have made him “comfortable” adjusting the “Follow the Bouncing Ball” task in ways that other MMR teachers might not be, though he also noted that ODE was “shooting for consistency” with the tasks because they would typically have collected data on the course implementation if not for the COVID-19 pandemic.

Table 23

Teachers Are Faithful to the Provided MMR Lesson Plans

Teacher	Supporting Interview Quotation
Adam	If they [the MMR lesson plans] suggest any kind of things that I should focus on before the lesson, any videos that they think I should watch to kind of help me strengthen the lesson or certain things that I need to do on certain days, I need to make sure that I'm doing those.
Beth	It wasn't much to plan; I feel like I hardly planned it at all. But they [ODE] really did a lot of the background work.
Cathy	With MMR, it has been a matter of pulling up the PDF of the lesson plan and looking at what they [ODE] suggest and I try to stay true to it, whatever they want the lesson to look like.
Debbie	I think that I can get them [the students] to identify rebound height through discussion, but they don't need all those words there.
Ethan	I look at the lesson plan and I look at the different objectives and what the tasks are doing. And I try and recall what we did last year and see if there's any changes.
Fred	I was going to rewrite that part [Part 2 of “Follow the Bouncing Ball”]. I don't know if the state [ODE] really wants me to do that. They're shooting for consistency.

Teachers Anticipate Challenges and Student Responses. The 8 MMR teachers anticipated various challenges as they planned their instruction, shown in Table 24. Beth, Debbie, Fred, and Gwen expected that their students might struggle with the mathematics content to some degree: Beth's concern was with interpreting decimal fractions (e.g., 1.5) as a "rise over run" when thinking about slope; Debbie thought that her students would need support to develop an understanding of what their linear regression models represented how they could be used to make predictions; Fred considered spending additional time discussing the meaning of a functional relationship; and Gwen anticipated difficulty recalling the Pythagorean Theorem and prior knowledge of similar triangles. Several teachers anticipated potential student responses and how they might react: Beth, Ethan, and Isabel were confident in their ability to ask questions that might help lead students toward solving problems on their own. Adam, Beth, and Ethan specifically voiced their concern with maintaining student engagement in remote settings: Adam's worry was helping his remote students visualize his classroom space, whereas Beth and Ethan foresaw issues with maintaining students' interest and engagement when learning online.

Table 24

Teachers Anticipate Challenges and Student Responses

Teacher	Supporting Interview Quotation
Adam	This assignment has to deal with something that's in the classroom that they're not physically in. So how can I help them see the classroom and apply it to their parts so that they can actually really be engaged in it?
Beth	They [students] will think of three halves, for example, as 1.5 instead of the "rise over run." Well, for that one I would just say, "What does a slope of three halves mean?"
Cathy	I were the students, other than measuring the size of your hand… what could possibly be relevant?... I'm not sure that they'll come up with more than one way to measure it.
Debbie	Where they [the students] will struggle is just, "What does this regression mean? I'm clicking a button and this equation is appearing. What is this?" I think it's just going to take some practice in understanding what that's really representing.
Ethan	Now that we're getting into more "mathy" stuff, stuff that they're struggling with, a little bit more stuff that's harder to do in an online environment, it's tough to keep them [the students] engaged and excited about what they're doing.
Fred	I anticipate having to discuss what it means, a functional relationship, that whole notion… I'm using sort of high vocabulary, but I want to make sure that I say it's if I do something, something else happens.
Gwen	They [the students] will probably have little-to-no recollection of Pythagorean Theorem or similar triangles.
Henry	I think about what questions I want to ask to get the students thinking…but I also think about what questions I expect from the students and how I'm going to reply to those questions.

Teachers Adjust for Face-to-Face, Remote, and Hybrid Learning Formats. As shown in Table 25, 6 of the 9 teachers discussed their adjustments and considerations for teaching in various learning environments (e.g., face-to-face, remote, and hybrid formats). The 2020–2021 year challenged teachers to establish safe learning conditions for students learning face-to-face with the spread of COVID-19, including the use of face masks, social distancing, and enhanced sanitation of school supplies and materials. Teachers also faced the numerous adjustments to online learning and balancing between groups of face-to-face and remote learners. These unprecedented obstacles required teachers to plan their instruction and their use of tasks accordingly to promote student learning.

Of the hybrid teachers, Cathy, Debbie, and Fred discussed ways in which they modified their tasks to make them safe for face-to-face learning. Cathy and Debbie took precautions with data collection involving Starburst™ candies and bouncing balls to prevent germs from spreading. Fred used instructional videos so that his students could watch balls bouncing in slow motion and record measurements. These three teachers planned for their students to work individually, rather than in groups as suggested in the MMR lesson plans so that their students would be less likely to spread germs. Debbie used Padlet as an alternative to sticky notes for her students to gather and organize their thoughts.

Adam, Beth, Cathy, and Ethan adjusted their tasks to be suitable for remote learning: Beth and Ethan adapted the MMR handouts for each of their tasks using technology such as Pear Deck and Google Slides. Both technologies allow teachers to view student work in real-time so they teachers can facilitate and provide feedback when appropriate. Adam discussed his desire to make the necessary tools and resources available to his online students, whereas Cathy explained her difficulty with making data collection “meaningful” because her online students would not be able to grab candy from the bowls that were used in class, so she designed an additional handout with questions for them to explore.

Table 25

Teachers Adjust for Face-to-Face, Remote, and Hybrid Learning Formats

Teacher	Supporting Interview Quotation
Adam	Is this something that they [the students] literally can't do without the supplies? So, I need to make sure that I could make it available for the students that are at home.
Beth	With Peardeck, I can see exactly what they [the students] are doing in real time.
Cathy	The one thing I don't like about this one [task]… I really just can't make the data gathering piece for the kids at home meaningful.
Debbie	Part of the rationale for that [Padlet] was the old-fashioned way [pre-COVID] might be just making a list on the board and or maybe just taking post-its, but they [the students] can't grab the posts and reorganize them because they can't share stuff.
Ethan	How could I take something that is supposed to be hands-on and make it into a virtual task? That's been my biggest challenge this year.
Fred	They [ODE] have actually been asking for ways people have been changing it [the task] to help it be virtual and also be safe with COVID and all those things… that it's social distancing for COVID and also making it accessible for online people.

Summary of Phase 2: Task Planning Findings. The MMR teachers planned their instruction according to the lesson plans that they had been provided by ODE, including the recommended instructional practices and pedagogy (i.e., having students brainstorm ideas, facilitating whole-class discussions, and asking questions rather than providing information). However, teachers with more experience with the MMR course or with a particular task tended to adapt and modify their tasks more so than others. Much of teachers' planning involved anticipating how a task might play out in the classroom and what instructional moves to make to support students' learning. The fall of 2020 presented teachers with new obstacles to overcome because of the COVID-19 pandemic, as many teachers had to find ways to implement their tasks safely in face-to-face settings and in remote settings. As they might in a typical setting, they also considered how their students might respond to a task and how they could answer students' questions or provide support to those who might struggle. Throughout the process, teachers were flexible and allowed their students to guide the length and direction of a task to some degree.

The Emergent Themes for Task Setup

The emergent themes for task setup are presented and described in Table 26. These themes answer the third part of research question 1, referring to teachers' task setup and reasons for doing so: (a) teachers ask, "What do you notice?" and "What do you wonder?" (b) teachers facilitate whole-class discussions before student work time; and (c) teachers communicate their expectations. These themes are discussed in the following paragraphs.

Table 26

The Emergent Themes for Task Setup

Theme	Description
Teachers ask, "What do you notice?" and "What do you wonder?"	• Students made observations and posed questions about real-world scenarios. • Students brainstormed ideas.
Teachers facilitate whole-class discussions before student work time.	• Teachers led discussions involving real-world scenarios. • Students identified independent and dependent variables to explore.
Teachers communicate their expectations.	• Teachers communicated classroom norms and expected behaviors. • Teachers set the tone for students' engagement with tasks.

Teachers Ask, "What Do You Notice?" and "What Do You Wonder?" Of the 8 MMR teachers, 6 of their tasks incorporated some aspect of the "What do you notice? What do you wonder?" routine, though not all of them used this exact language. Those who did not, however, included a similar brainstorming phase within task setup. Shown in Table 27, 5 of these teachers discussed this aspect of task setup during their interviews; Fred used the same task as Debbie and Ethan and included a brainstorming intro but did not provide a supporting interview quotation to document here.

Table 27

Teachers Ask, "What do you Notice?" and "What do you Wonder?"

Teacher	Supporting Interview Quotation
Adam	When we came back, I'm like, "So did you watch the video? Let's go through it, what questions do you have? What did you notice? What did you wonder?"
Cathy	I like that we start with the "What do you notice? What do you wonder?" quite a lot. I've been using that for the last couple of years.
Debbie	Towards of beginning of a unit, I kind of like that to be more of an exploration stage and I like the "notice and wonder" type of thing.
Ethan	Jamboard works pretty well… They [the students] were really good… at putting posts up and saying, "What do you notice? What do you wonder?"
Gwen	The task starts with making observations, what they [the students] notice and wonder about a ramp, and making decisions on what they can do with that information.

Debbie, Ethan, and Fred introduced their task with a video of the 1965 Super Ball commercial. Rather than using a video, Cathy's students made conjectures about a bowl of candy that she presented at the front of the classroom whereas Gwen took her students to the location of a ramp in their school building. Adam's students discussed what they noticed and what they wondered about their physical classroom environment, though it was not observed by the researchers. In each case, the video or object served as a means of generating student discussion focusing on the context of each task. Such discussions transitioned into conversations about relevant independent and dependent variables that might be explored in each task, extending from and including students' ideas.

Beth, Henry, and Isabel did not include a preliminary brainstorming session. Though Beth and Henry are familiar with the "What do you notice? What do you wonder?" routine through their experience teaching MMR, both teachers set their students to work on their respective tasks without conducting a full-class discussion. Neither explained their reasoning for this decision, but perhaps it was because the MMR lesson plans did not call for such discussions before each task, or because neither task centered around a real-world context. Alternatively, the lesson plans for the other tasks included a brainstorming segment and each task involved a real-world project or experiment.

Teachers Facilitate Whole-Class Discussions Before Student Work Time. The 8 MMR teachers regularly facilitate whole-class discussions before allowing their students to engage with tasks independently or in groups. Table 28 provides interview quotations for each of the 6 teachers whose observed tasks involved an initial whole-class discussion. Though Beth and Henry's observed tasks did not include a preliminary discussion, both teachers explained how this was a common practice; Beth's tasks "always have a discussion point," she said.

Table 28

Teachers Facilitate Whole-Class Discussions Before Student Work Time

Teacher	Supporting Interview Quotation
Adam	We talked about surface area and how to find it for a rectangle. I had them [the students] guess and do like a high, low type of thing: what's the absolute low that this can be? What's the absolute high? And then put your educated guess in the middle.
Cathy	I liked having the ability, even though I guided it more than I should have, to go back then through their list and say this one we'll control. This one is going to be constant.
Debbie	It helps to have that little video clip first with the Super Ball. They [the students] were really interested in that.
Ethan	They [the students] did do a good job of going through the “What do you notice? What do you wonder?” And eventually we rounded it out towards, “What attributes can you measure, and which ones do we think…” so like guiding them back to, “Let's talk about drop height versus rebound height.”
Fred	I thought that at first, the discussion went pretty well. I thought what they [the students] were suggesting as, what would affect the bounce and what would not affect the bounce, were right on.
Gwen	Researcher: did you plan to ask those questions about ramps? Gwen: I did, because I think it's important to know internally what the students know or where their preconceptions are before you dive into something like that.

The 6 teachers included in Table 28 led discussions focusing on (a) mathematics content, (b) variables of interest to explore, and (c) the real-world components of the tasks. Regarding mathematics content, Adam’s class “talked about surface area and how to find it for a rectangle," whereas Beth opened a discussion by asking her students what they remembered about slope. Cathy, Debbie, Ethan, and Fred's tasks involved collecting data; therefore, their initial discussions focused on determining appropriate variables and how they could be measured. In terms of nonmathematical discussions, Adam and his students discussed the dropped ceiling in their classroom, whereas other classes focused on 1965 Super Balls and ramps in real life.

Teachers Communicate Their Expectations. The four teachers displayed in Table 29 communicated their expectations for students before they engaged with the tasks. Though Beth's task did not have a pronounced setup phase, she frequently provides the expectation for her students to communicate their answers verbally, similar to the expectation that Gwen set for her students during task setup. In addition, Beth regularly requires that all her students participate actively with their group members and collaborate. Adam required his students to thoroughly document the evidence that they incorporated into their group presentations, including calculations, diagrams, and costs associated with various aspects of their remodeled classrooms. He placed particular emphasis on the task rubric that was provided with the MMR materials, referencing students back to it when they had questions and needed additional clarification. Because Isabel's task differed from what her students were accustomed to, she briefly explained her desire for correct answers and for her students to check in with her regularly throughout their engagement with the task.

Table 29

Teachers Communicate Their Expectations

Teacher	Supporting Interview Quotation
Adam	Day 1 is always about introducing the handout… Giving them the rubric and going through the rubric in detail, letting them know exactly, "These are the markers, and this is what's expected of you."
Beth	On the Google slides, I had just reminders, my expectations… the expectations that all students are participating… Everybody's speaking, everybody's working, they're on the same problem at the same time.
Gwen	Researcher: You said for the students to verbalize how they approached each problem as they were working. Is that one of your expectations for them? Gwen: Yes. I can't fully grasp what they're thinking by just looking at what they have on the paper. They need to be able to, using that word verbalize, not just to read what they have but to… fully show and represent that they understand it.
Isabel	I explained to them [the students] how it's just an activity… I'm expecting them, though, to do the activity correctly. There are right and wrong answers, and I don't let them turn it in with wrong answers.

Summary of Phase 3: Task Setup Findings. The MMR teachers in particular set up their tasks according to the lesson plans provided by ODE. They included the suggested brainstorming segments and facilitated whole-class discussions about both the real-world contexts and the mathematics involved with each task, specifically focusing on determining variables that students might explore. Many of the teachers set expectations for their students' engagement with the tasks, such as expectations for group work and collaboration and for written and verbal reasoning.

The Emergent Themes for Task Implementation

The emergent themes for task implementation are presented and described in Table 30. These themes answer the last part of research question 1, referring to teachers' task implementation and reasons for doing so: (a) teachers encourage productive struggle, (b) teachers elicit evidence of students' thinking and reasoning, (c) teachers monitor and facilitate student engagement, (d) teachers ask questions, and (e) teachers provide instructional support. The following paragraphs describe how the teachers demonstrate each theme.

Teachers Encourage Productive Struggle. As shown in Table 31, the teachers in the present study expressed their belief in *productive struggle*, that is, challenging their students solve mathematical problems without providing direct guidance and worked-out examples during instruction. Debbie acknowledged the "danger" of "doing all the thinking for the student," recognizing that students learn more mathematics when they think for themselves. Adam and Beth noticed shifts in their instructional practices that they attributed to teaching the MMR course, realizing that they allowed their students more time to think on their own without intervening. Ethan and Gwen recognized that their tasks suggested pathways for students to follow but did not do the mathematical thinking and interpretation for them. Ethan facilitated his students' engagement as they collected data and determined linear regression models but asked questions and prompted his students to make sense of the models they generated. Henry's students directed the mathematical discussions that took place in class as they solved graphing problems using Desmos, explaining their work and demonstrating it on the smartboard as he asked questions to guide their thinking.

Table 30

The Emergent Themes for Task Implementation

Theme	Description
Teachers encourage productive struggle.	• Teachers do not provide step-by-step procedures or direct guidance. • Teachers encourage students that mistakes are part of learning.
Teachers elicit evidence of students' thinking and reasoning.	• Teachers require students to explain their answers verbally or in writing. • Teachers require documentation and evidence from data.
Teachers monitor and facilitate student engagement.	• Teachers listen to students' conversations and provide comments. • Teachers intervene when appropriate.
Teachers ask questions.	• Teachers ask questions to direct students' thinking. • Teachers ask questions to shift the thinking process back to the students.
Teachers provide instructional support.	• Teachers help students visualize aspects of their tasks. • Teachers prompt students to reevaluate their thinking.

Table 31

Teachers Encourage Productive Struggle

Teacher	Supporting Interview Quotation
Adam	I'm finding myself not jumping in and saving them [the students] like I typically used to do… I want them to think something through and then give me an answer.
Beth	I liked that they [the students] were willing to do some productive struggle… They were really tolerant of me just asking them more questions and they didn't give up.
Cathy	I would like them [the students] to struggle with that [measuring and collecting data] a little bit, not give it away too much. That's the hardest thing, is to not give away too much. Let them struggle.
Debbie	I think there's a danger of, you do all the thinking for the student and then they don't have the opportunity to struggle with it a little on their own… that's a skill that we're really trying to build.
Ethan	Trying to figure out the best way to get them [the students] back to that point without just telling them the answer, I think that's one of the best things that good teachers do well.
Fred	I didn't teach them [the students] "how to do it" right before I gave it [the task] to them. I was giving them "just in time" help and asking them what they've done with it.
Gwen	It's important for me just to reiterate that they [the students] are not going to be learning through direct instruction. And to me, I think with this task, that was a really big goal.
Henry	I think it helps where they [the students] can make mistakes. And that's part of the whole process for this class, I think, is getting them comfortable to make mistakes.
Isabel	I keep trying to keep that [decline in cognitive demand] from happening… and not giving them [the students] too much.

An aspect of productive struggle is acknowledging that mistakes are part of the learning process. Adam, Ethan, and Henry embraced this philosophy, providing learning environments where mistakes were acceptable. They encouraged their students to try and praised them for their contributions. “I want them [the students] to be able to take a risk,” Adam explained. Similarly, Ethan acknowledged that “Failure doesn’t mean the end of the world, it just means you have to try again.” Henry encouraged one of his students who was hesitant to type an expression into Desmos that there was no penalty for incorrect attempts, that he could always “just erase it and try something else” if his first try was unsuccessful. “They can make mistakes, and that’s part of the whole process for this class [MMR],” Henry concluded.

Though many of the teachers challenged their students to grapple with mathematical concepts and ideas, not all their students were productive in doing so. Evidence from IQA scores shows that Adam, Cathy, and Gwen’s students did not engage students in high-level mathematical thinking and reasoning. Moreover, Cathy and Gwen seemed to notice this to some degree; Gwen acknowledged her students’ difficulty getting started, struggling to the point of being unproductive: “The two [students] that understand more from the beginning go. The other two just kind of stared… and kind of ignored me. They completely shut down.” Cathy felt that her students lacked the need to struggle with her task: “I don’t think they [the students] had enough to struggle, really… I don’t think there was much struggle left.” She attributed this, in part, to having a “good group” of students who had become familiar with the process of defining variables, collecting data, and interpreting the results.

Teachers Elicit Evidence of Students' Thinking and Reasoning. The 8 MMR teachers communicated a desire for their students to explain their thinking and reasoning and made instructional moves encouraging them to do so, as shown in Table 32. Teachers such as Beth, Gwen, and Henry emphasized their desire for students to communicate their thinking and their work verbally, whereas Debbie's focus was on the use of data as evidence and Adam required documentation in the form of equations, costs, and even websites referenced. Beth explained that, for students, communicating their reasoning verbally is "much more difficult than just showing somebody" and "makes it [mathematics] more in depth than what they [the students] are used to doing." Likewise, Henry feels that his students "aren't there yet, if they can't explain it." Debbie provides as many opportunities as possible for her students to explain trends using data because "they don't have a lot of skills in those areas" when they reach her class.

Adam expects students in all his classes to "show their work" because it helps him to evaluate their level of understanding. "If it [written work] shows me a pattern or a level of how you're processing a question, that's beneficial to me as your educator," he explained. Similarly, Ethan expressed that student explanations help him to "understand why students answer the way they do," which then guides his own thinking and future instruction. Ethan and Henry also believe that students can enhance each other's mathematical understanding when they make their thinking visible to the whole class. To Ethan, it helps other students to realize what their classmate are thinking. For Henry, students come to the realization, "Oh, I hadn't thought about that," and become aware of ideas that might also "trigger" something in their own minds.

Table 32

Teachers Elicit Evidence of Students' Thinking and Reasoning

Teacher	Supporting Interview Quotation
Adam	Everything had to be detailed. You had to use a spreadsheet, had to show formulas. You had to show all of your logic and reasoning and you had to show data, pictures.
Beth	I'm trying to get them to explain it [students' work] in words now…They have to think about how they can explain it using the right vocabulary that someone else could understand… It definitely makes it more in depth than what they're used to doing.
Cathy	I think it's describing their [students'] work that that is the biggest challenge for most of them still… it's a great skill to have. So, I'm glad we're pushing it with them.
Debbie	I'm kind of a stickler about this number 14, where it says use evidence from class data… they [the students] need to tell me specific things that they see from the data.
Ethan	I always felt like it's great to hear somebody explain their answer or say what their answer is.
Fred	Someone said, "It's where the graph starts…" So, what do you mean by "starts" there? That's what I was trying to get at.
Gwen	I can't fully grasp what they [students] are thinking by just looking at what they have on the paper… It's more about being able to reproduce something. You have to be able to explain what's going on as well.
Henry	I think having them explain makes them think more about it [mathematics] and then it makes them have some deeper thinking about what they're doing.

Adam stands out from the other teachers in terms of the evidence he expects his students to provide. With the "Remodeling Our Classroom" task, he prompted his students to provide evidence for their mathematical work, such as the calculations and conversions they used to determine the cost of paint to cover the classroom walls, when assigning the task (i.e., during task setup). During implementation, however, he placed significant emphasis on *nonmathematical* evidence, including accurate documentation for objects that students included in their remodeled classrooms (e.g., screenshots from the internet) and labor costs for removing chalkboards from the walls. His students were also expected to document the online resources and websites that they used to develop their proposals. Though Adam set the expectation for his students to provide *mathematical* evidence of their thinking and reasoning in their budget proposals, he did not follow through as his students worked in groups. Instead, he remained silent and rarely questioned his students as they worked.

Teachers Monitor and Facilitate Students' Engagement. As shown in Table 33, several of the teachers described their endeavor to monitor and facilitate their students' engagement with the tasks. Adam, Beth, Ethan, and Henry expressed a similar pattern of closely observing their students' group conversations and listening for instances where they struggled with a particular idea. Adam, Ethan, and Henry generally follow the same approach, allowing their students to progress on their own or in groups unless they see or hear something that causes them to pause; at this point, they either ask a question, provide a suggestion, or bring the whole class together to discuss the issue.

Table 33

Teachers Monitor and Facilitate Students' Engagement

Teacher	Supporting Interview Quotation
Adam	A lot of times I'm just simply observing. Sitting back, listening to the conversation and listening to what they're coming up with and possibly answering questions.
Beth	Once I saw that one group had graphed the point backwards, I started actively checking for that.
Cathy	They [the students] brainstormed some of the things that might impact the grabs. And then I just was flipping through their documents and writing down what they wrote… I get to see everybody's and they're all accountable.
Ethan	I'm going to bounce from group to group, and if they're all moving in the right direction, I'll keep them going.
Gwen	I serve as a facilitator in all my classes, not just MMR.
Henry	I'll walk around and watch the kids and… if I see a pair that's really struggling, then we'll come back as a class and have other people bounce ideas around.

For example, Beth noticed that some of her students were graphing ordered pairs (x, y) "backwards," as if y was the horizontal axis and x was the vertical axis. After she detected and addressed this misconception, she began checking for this issue as she relocated from one breakout room to the next. Cathy described how she would typically walk around the classroom and observe students' work and listen to their conversations but could not this year because of the COVID-19 pandemic. Ethan teaches remotely but explained that his ability to facilitate groupwork is enhanced when he can listen to multiple groups from across the classroom and move to where he is needed. Without being in all the breakout rooms at once, it is more difficult for him to determine where students' conversations are at any given point in time.

Gwen was the only teacher to explain the practice of taking notes during class, recording her reflective observations as her students worked on problems in groups or individually. Gwen explained that taking notes was especially "useful" when she wants her students to present or share their work. "Student A might be at this point and Student B might be here; and you want to scaffold the order in which they're sharing… not just whoever wants to share." She believes that the order in which students share their work matters and that it is important for not only "the smart kid" to share. However, she did not express the particular ways in which she structures whole-class debriefs during the limited timeframe of the interviews.

Teachers Ask Questions. The 9 teachers asked a plethora of questions to their students and did so for various reasons, as depicted in Table 34. Three such reasons are common among many of the teachers: the first is to scaffold or guide students' thinking from one idea to the next. Henry anticipated that if his students struggled with graphing appropriate inequalities, he might ask questions such as, "Does this have any connections with what you did yesterday?" to direct students back to what they had done previously and make connections between the two. Fred asked his students which variables they "have control of" to help them determine between independent and dependent variables, hoping that they would realize the distinction between the two types.

Table 34

Teachers Ask Questions

Teacher	Supporting Interview Quotation
Adam	I'm constantly thinking, "Okay, I know this question is going to come. What question can I immediately get back to them [the students] to get them to try to think about it?
Beth	There were a couple of groups that said, "I'm bad at slope. Can you help us?" And I said, "Okay, well, what do you remember?"
Cathy	I want student engagement; and that's the way you get student engagement, is you ask a question.
Debbie	If it's something I'm really wanting them [the students] to try to understand, I try to respond with a question if I can, just because the thinking stops if I just answer.
Ethan	It's always like, how do I get them [the students] there without telling them the answer. Asking them leading questions to try and get them to take baby steps, to then take that big leap forward.
Fred	Often in my teaching, even if they [the students]—especially if they—give a correct answer, I actually ask them, are they sure? I'm trying to deliberately break the mold. Teachers have usually asked, "are you sure?" when the student's got it wrong.
Gwen	As exhausting as it is to ask question after question, I don't think I would change that aspect of it [task implementation]. I don't want to be the one that gives them all of the information.
Henry	A good class is, I ask a question and just let the students have at it. And they can ask questions, but they know 90 percent of the time I answer their question with another question. I try to avoid giving answers.
Isabel	A question makes them [the students] think, whereas a response just makes me give them an answer and they, whether they understand it or not, "okay, let's move on, because now I have the answer."

The second reason for teachers' questions is to maintain the thinking process with the students rather than shifting it to themselves: "I don't want to be the one that gives them all of the information," Gwen stated. Debbie and Isabel shared a similar viewpoint that the thinking stops after they answer a question, whereas students continue thinking if they ask a question instead. Third, some of the teachers ask questions to assess students' understanding: when asked for his reasoning, Henry replied, "I do it [questioning] to assess the student." During our second interview, Ethan recalled asking a student, "What do you mean?" so that he could verify how the student determined the rebound height of a ball using their linear model.

Teachers Provide Instructional Support. The teachers provided instructional support for a variety of reasons, as shown in Table 35. To guide her students toward exploring drop height and rebound height, Debbie picked up one of the balls she had brought to class and started bouncing it so that her students would come up with the desired outcome on their own. In their follow-up discussion of their linear models, Fred purposefully rewrote the decimal 0.6 as 6/10 so that his students would recognize it as a rate of change, a change in y (rebound height) divided by a change in x (drop height). Beth encouraged students to "draw a picture" when they struggled to make progress with the "Discovering Slope" task. Gwen did the same when her students had difficulty visualizing what the proportions 1:12, 1:16, and 1:20 represented in the "Ramp It Up" (King, 2015) task.

Table 35

Teachers Provide Instructional Support

Teacher	Supporting Interview Quotation
Adam	If I hear someone just trying to take the low road to something… I'll unmute my mic and say, "okay, so let's think about what you just said." It's like a hint, and all of a sudden, they're like… "maybe there's something more that I need to do for this."
Beth	That group definitely wanted to use the formula. The problem is they don't have two points and they didn't really recognize that. That's why I kept trying to get them to go back to the "rise over run."
Debbie	I could see that they [the students] were going all these different directions. I thought, "they need a visual…" When they saw that [the ball bouncing], it seemed like it helped them.
Ethan	I just repeated what he was saying… "So, you're putting this in for x, which is the drop height." And he goes, "Oh, no." So then, he switched it back and he was like, "No, because that's the rebound height."
Fred	They [the students] say "rise over run," but then they see this decimal [0.6]. And so, me writing it as 6/10, I think was important to connect that this is still a slope.
Gwen	They [the students] didn't really understand what the ADA specifications were referring to, what they meant… So, at that point I had to intervene, otherwise, they would just keep going and just completely get lost.
Isabel	The first time, whenever they [the students] get anything wrong, I just put a little star next to it and say, "these are the things that are wrong. You need to go back and try and fix them…" A lot of times, that's enough to then get them to fix things.

Summary of Phase 4: Task Implementation Findings. The MMR PD appears to have also influenced the teachers' task implementation, as many of the themes relate to the pedagogical strategies encouraged in the training. Though the teachers provided instructional support in various ways, they typically did not remove the challenge for students by providing students with exact procedures and approaches to solve the tasks (except in instances where the cognitive demand of the task declined). Instead, they prompted students to attempt problem-solving strategies such as drawing pictures or diagrams and frequently asked questions to guide students' thinking. Specifically, teachers such as Ethan and Adam explained their desire to support productive struggle and encourage students that making mistakes is part of the learning process in MMR. Monitoring and facilitating students' engagement with the tasks, as well as eliciting evidence of students' thinking and reasoning, are aspects of teaching the MMR tasks that were emphasized in the PD but also contributed to the high-level cognitive demand of many of the tasks.

Task Analysis Across Cases

This section is divided into three parts: (a) trends in IQA scores across task phases, (b) trends in teachers' TAG classifications across task phases, and (c) mismatches between IQA scores and TAG classifications. The first part of this section concerns IQA scores so that comparisons can be made between these and teachers' self-reported TAG classifications. The second part addresses the second research question, focusing on teachers' perceived changes in the cognitive demand of their tasks and reasons for such change. The third part of this section focuses on mismatches between IQA and TAG analyses, addressing the third research question. Isabel's data are excluded in portions of this section because of the lack of Phase 3: Setup and Phase 4: Implementation data. Similarly, the following sections include comparisons among Phases 1–4 only; Setup 1: MR, Setup 2: CF, and Implementation: Student Work scores are not included in cross-case analysis because of gaps in available data.

Trends in IQA Scores Across Task Phases

The 8 MMR teachers' tasks (not including Isabel, who did not teach MMR) changed dynamically throughout the four phases of selecting, planning, setting up, and implementing. As shown in Table 3 on page 162, 3 teachers' tasks maintained their level of cognitive demand from Phase 1: Selection to Phase 4: Implementation: Fred and Debbie, whose tasks scored a 4 on the IQA rubrics across the four phases, and Beth, whose task scored a 3 consistently. Four teachers' tasks declined in cognitive demand from Phase 1: Selection to Phase 4: Implementation: these teachers are Adam (decline from 4 to 2), Ethan (decline from 4 to 3), Gwen (decline from 3 to 2), and Cathy (decline from 3 to 2). Henry's task surprisingly increased in cognitive demand from IQA scores of 3 in Phases 1–3 to a score of 4 in Phase 4: Implementation. A visual representation summarizing this information is provided in Figure 6.

In Figure 6, the Phase 1: Selection–4 IQA scores are represented as vertical bars, one for each task phase. The same four colors are used to signify each teacher's task at each of the four task phases: blue represents Phase 1: Selection, green represents Phase 2: Planning, yellow represents Phase 3: Setup, and orange represents Phase 4: Implementation. The vertical axis represents IQA scores, whole numbers 0–4, and each participating teacher is displayed along the horizontal axis. The horizontal dotted line segment indicates the separation between tasks that are considered high in cognitive demand (i.e., IQA scores of 3–4) and those that are considered low in cognitive demand (i.e., IQA scores of 0–2). Using this representation, readers can track the progression of each teacher's task from Phase 1: Selection to Phase 4: Implementation. For example, Cathy's task scored a 3 for Phases 1–3 but scored a 2 for Phase 4: Implementation. As shown in Table 3 on page 162, Adam's task was not scored for Phase 3: Setup, and neither was Isabel's for Phases 3–4.

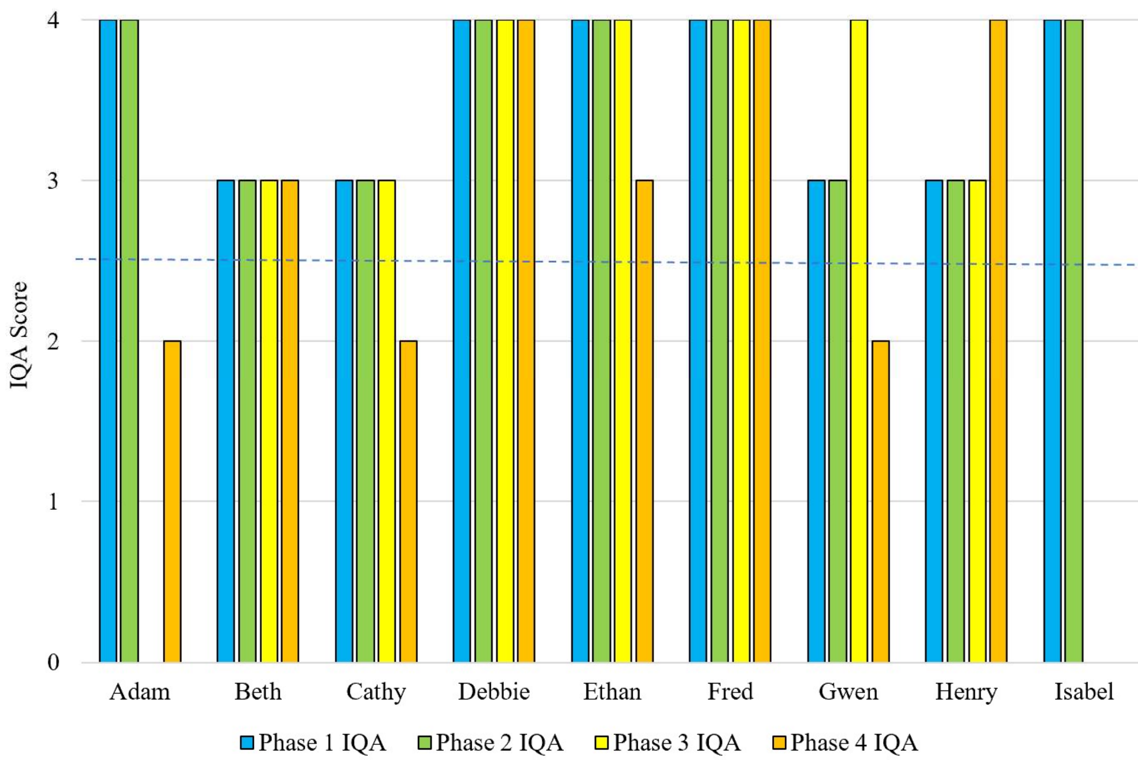
In the following paragraphs, I elaborate on 5 trends based on the initial Phase 1: Selection IQA scores and final Phase 4: Implementation scores of the teachers' tasks:

- the two tasks that initially scored a 4 on the IQA and maintained cognitive demand (Debbie and Fred's),
- the two tasks that initially scored a 4 on the IQA but declined in cognitive demand (Adam and Ethan's),
- the task that initially scored a 3 on the IQA and maintained cognitive demand (Beth's),
- the two tasks that initially scored a 3 on the IQA and declined in cognitive demand (Cathy and Gwen's), and
- the task that initially scored a 3 on the IQA but inclined in cognitive demand (Henry's).

Then, I discuss another common trend, that all the MMR teachers' tasks were high in cognitive demand and maintained identical IQA scores from Phase 1: Selection to Phase 2: Planning.

Figure 6

IQA Scores Across the Phases of the Teachers' Tasks



Note. The dashed horizontal segment separates high cognitive demand tasks (IQA scores of 3–4) from low cognitive demand tasks (IQA scores of 1–2).

The Tasks That Maintained an IQA Score of 4. Debbie and Fred used the same "Follow the Bouncing Ball" task provided in the MMR course materials. Therefore, the

Phase 1: Selection IQA scores for both teachers' tasks were identical. Similarly, both teachers made relatively few changes to the task, resulting in identical scores of 4 for Phase 2: Planning. Debbie and Fred's modifications were minor at this phase, but different, despite earning identical IQA scores: Unlike Fred, Debbie adjusted the format of the MMR task handout by separating large paragraphs into smaller groups of sentences. Each teacher added a unique component to the task handout, as described previously: Debbie included a table to help students organize the data for their experiment, whereas Fred provided specific instructions for students to watch videos. Moreover, each teacher's choice of technology for the task was different. Fred's students used Excel spreadsheets, but Debbie chose to use a more user-friendly software that required less instruction and guidance. Recall also that Fred added the question "What does the slope of your best-fit line mean in the context of this data?" to the task handout. Though each teacher's planning for the task was different, neither resulted in a decline in cognitive demand from an IQA score of 4 because their adaptations did not negatively impact the mathematical work and thinking expected of students.

Although Debbie and Fred's setups were unique, they contained similar elements that resulted in maintaining the high cognitive demand of the task. Both teachers engaged their students in exploring potential variables of interest, including drop height and rebound height. During this common segment of instruction, Debbie and Fred elicited students' ideas instead of providing information themselves. Moreover, both teachers led discussions of potential independent and dependent variables, focusing on the meaning of each type of variable and how the former influences the latter. Students in both classes successfully identified that the drop height of a ball influences its rebound height, suggesting that the former is the independent variable, and the latter is the dependent variable.

Task implementation in both instances included student work time, followed by a whole-class discussion in which students voiced their findings and the conclusions they had drawn. Both Debbie and Fred asked questions prompting students to explain the linear relationship among their data, resulting in IQA scores of 4 because students indeed provided such evidence of their thinking and reasoning. Though Fred had to press his students more so than Debbie to get them to interpret the linear trend in their data, his students reached the conclusion that, "For every 10 cm on the drop height, the rebound height increases by 6 cm." This statement that was not expressed the same way in Debbie's class, but her students still identified and explained the linear pattern in their data. Both teachers' tasks scored 4 on the IQA for implementation, but the precise language used by students and the amount of prompting done by the teacher was slightly different in each case.

The Tasks That Declined From an IQA Score of 4. Recall that Ethan's task was the same as both Debbie's and Fred's and therefore scored a 4 on the IQA for Phase 1: Selection. Unlike Debbie and Fred, however, Ethan's district required him to teach remotely at the time of the observations; this led him to adapt the format of the handout from a Word or PDF document into a series of Google Slides that his students could manipulate in their online learning groups. Though he changed the format of the handout, the prompts for students and the mathematical work required of them remained relatively unchanged, resulting in a similar IQA score of 4 for Phase 2: Planning. This suggests that the same Phase 1: Selection task may yield identical IQA scores for Phase 2: Planning, despite differences in teachers' type of learning environment (face-to-face, remote, or hybrid). However, it is certainly feasible that, for example, Ethan may have lowered the cognitive demand of the task in Phase 2: Planning, and possible that the cognitive demand lowered *because of* adaptations that he made for his remote learning environment. However, this was not the case in the present study.

The Phase 3: Setup score for Ethan's task remained a 4, like Debbie and Fred. Though Ethan himself suggested that his students should use rebound height as one of the variables to explore, this was already provided for students on the original MMR handout for the task. Therefore, the IQA score did not lower from a 4 to a 3. The remainder of the task setup progressed in a similar fashion to both Debbie's and Fred's, including the two components: (a) a discussion of the real-world context, including the 1965 Super Ball television commercial; and (b) a discussion of the variables of interest and the relationship between dependent and independent variables. The setup phase of instruction in Ethan's remote environment progressed in a similar way to the two that occurred in face-to-face learning environments, though some differences occurred.

The implementation of Ethan's task differed more so from Debbie and Fred's and scored a 3 on the IQA for several crucial reasons: one was that Ethan provided more structure and guided some aspects of the task that Debbie and Fred facilitated, but let their students do on their own. For example, Ethan shared his screen with his students and led them through the process of creating a prediction graph for the relationship between the drop height and rebound height, as described earlier in the chapter. He created the graph, labeled the *x*- and *y*-axes, and provided several *x*-values (drop heights, in cm) for students to use, simply asking them to predict what the associated *y*-values (rebound heights, in cm) might be. Alternatively, Debbie and Fred's students thought through this process on their own, supported by their teachers as they walked around their classrooms and monitored students' progress.

Another noticeable difference was in the rigor of students' responses to Ethan's questions and the connections they made between the graphs of their data and the real-world context. Though Ethan questioned and pressed his students for evidence of their thinking to the same degree as Fred, Ethan's students did not provide the level of responses that Fred's students provided. Though it is not clear why, perhaps Ethan's students were either unwilling or unable to provide such responses. Debbie and Fred's students explained that the positive slopes in the graphs of their data represented increases in rebound height, corresponding to increases in drop height. Ethan's students noticed the positive slope and, with his support, attributed the greatest slope to the "best bouncer," but did not directly connect the slope back to the data and the real-world situation.

Though Adam's task was different than the other three described in the previous paragraphs, it contained similar elements that resulted in IQA scores of 4 for Phases 1–2. His implementation of the task, however, differed from Debbie and Fred's even more so than Ethan's, resulting in an IQA score of 2. One of the key differences was that Adam's students engaged in some form of problem-solving, but it was not necessarily mathematical in nature; many of the decisions Adam's students made involved real-world aspects, such as whether to remove objects from the walls or what additional objects to include in the remodeled classroom. Most of their mathematical work related to the procedures they used to calculate the cost of the project, including unit conversions and calculations of surface area. A second crucial difference was that Adam allowed his students to divide the work of the task among themselves so that each student in a group worked on a different piece. This might have been the reason for limited dialogue and communication between students during the observed lessons.

Third, Adam rarely elicited evidence of students' thinking and reasoning during task implementation, especially when compared to Debbie, Ethan, and Fred; these three teachers consistently questioned their students and asked them to contribute to in-class discussions. Adam explained that he frequently tended to listen in to his students' conversations without saying anything if he thought that they were making connections on their own. There were instances where he could have asked a strong question to help make his students' thinking visible and even extend their thinking; for example, Adam asked one student whether she planned to remove objects from the walls or leave them in place and adjust her calculations accordingly. Though the student stated that she planned to remove the objects from the walls, Adam could have asked her to explain her process for calculating the surface area of the classroom and how she knew that her approach was reasonable.

Finally, Adam guided his students through the process of determining the length of the pipeline in the classroom and provided examples demonstrating his expectations for the scale drawing that each group was expected to complete. In doing so, Adam provided more support beyond what Ethan did, removing a greater amount of the mathematical thinking required from his students to complete the task. These actions contributed to the decline in cognitive demand from an IQA score of 4 in Phase 2: Planning to a score of 2 in Phase 4: Implementation.

The Task That Maintained an IQA Score of 3. Of the 4 teachers whose tasks scored a 3 on the IQA for Phases 1–2, Beth was the only one whose task maintained its cognitive demand across the four phases. Like many of the other teachers involved in the present study, Beth made minimal adjustments to the Phase 1: Selection version of the “Discovering Slope” task, resulting in an identical Phase 2: Planning score of 3. Though Beth reformatted the handout into a Pear Deck presentation to use with her remote students, the content of the task remained nearly identical. The lack of a pronounced setup phase of instruction contributed to the maintenance of cognitive demand in Phase 3: Setup, as there were few opportunities for the cognitive demand to change in such a short span of time. Finally, the cognitive demand of the task was maintained during implementation, resulting in an IQA score of 3 for Phase 4: Implementation as well; though a couple students provided verbal explanations that might score as 4 on the IQA, most students’ work and communication aligned with the criteria for a score of 3. They engaged in problem-solving, executed mathematical procedures, and made some connections between various concepts and mathematical representations.

The Tasks That Declined From an IQA Score of 3. Beth, Cathy and Gwen made few changes to the tasks they chose for the present study, resulting in identical scores of 3 for Phase 2: Planning. Despite this, the three task setups were unique. Beth did not engage her students in a whole-class discussion, but Cathy and Gwen did. Moreover, Cathy's task scored a 3 for Phase 3: Setup whereas Gwen's task scored a 4. Both teachers engaged their students in discussions of the real-world context and mathematical relationships involved with their respective tasks. The primary difference between how Cathy and Gwen set up their tasks was that Gwen asked her students to verbalize how they approached each problem as they worked in their groups. By doing so, Gwen enhanced the cognitive demand of the task by adding a component that was not written on the handout; this instructional move resulted in an increase from an IQA score of 3 to a score of 4.

Both Cathy and Gwen's tasks scored a 2 for Phase 4: Implementation despite the high-level task potential in Phase 3: Setup. In both instances, the students failed to exhibit evidence of high-level thinking and reasoning and the task focused on procedures without mathematical connections. Both teachers provided support for their students, such as suggestions for how to collect and record data and how to visualize a problem they were solving. These instructional moves did not necessarily influence the decline in cognitive demand; instead, the reason for the decline of each task was different.

Cathy lowered the cognitive demand of the "Starbursts™ Grab" task by asking only procedural questions to her students, such as to evaluate a function at specific values. None of the questions she asked prompted students to make mathematical connections and provide evidence of their understanding, limiting the rigor of the task to the level of *procedures without connections*. Moreover, she eliminated the need for students to make such connections by making them herself: for example, she was the one to conclude that the linear model the class generated was not a strong fit to the data. Unlike Cathy, Gwen attempted to promote high-level thinking by asking questions and suggesting that students draw diagrams to help them visualize the mathematics they were doing. However, her students leaned on the procedural approach of solving proportional equations (e.g., $1/16 = 15/x$) and their explanations were limited to descriptions of the procedures they used. Both Cathy and Gwen's tasks scored a 2 on the IQA for Phase 4: Implementation, but the reason for the decline in cognitive demand was different in each case. Cathy's facilitation contributed most to the decline of her task, whereas Gwen's task declined in cognitive demand because of the approaches that her students took to complete it.

The Task That Inclined From an IQA Score of 3. Henry's was the only of the MMR tasks with a higher Phase 4: Implementation IQA score than Phase 1: Selection, indicating that it was the only task that was implemented at a higher level of cognitive demand than its original potential. Phases 1–3 of Henry's task were similar to other teachers in some regard: like Beth, Cathy, and Gwen, the Phase 1: Selection version of the task did not explicitly prompt students to provide evidence of their thinking and reasoning, yielding an IQA score of 3. Henry made no changes to the task prior to his instruction and did not engage his students in a preliminary discussion, comparable to Beth. However, Henry frequently asked his students to explain and present their solutions to the rest of the class; their explanations led to an increase in cognitive demand for Phase

4: Implementation. Like Gwen, he introduced an element to his task that was not present in the previous phases, though Gwen's addition occurred in Phase 3: Setup.

Identical, High IQA Scores Among Phases 1 and 2. As shown in Figure 6 on page 353, each of the 8 MMR teachers' tasks received the same IQA score for Phases 1 and 2. That is, none of their tasks changed in cognitive demand significantly enough to warrant a change in IQA score between Phases 1 and 2. In fact, 6 of the 8 tasks were scored identically in Phase 3: Setup as well, with only Gwen's task increasing from a score of 3 to 4 and Adam's task without a Phase 3: Setup score because of missing data. The Phase 1: Selection and 2 IQA scores for all the MMR tasks were either 3 or 4, indicating that the cognitive demand was high for every task in those phases.

These two results could be related to each other: perhaps one reason that the MMR teachers did not significantly modify the mathematical aspects of their tasks in Phase 2: Planning was because the cognitive demand was already high. Many of the teachers desired to maintain the original intent of the MMR materials, whether it be because of the research done by ODE or for other reasons. Though ODE did not collect data during the 2020–2021 academic year due to the COVID-19 pandemic, some teachers still voiced this reason for maintaining the fidelity of the course materials. Cathy explained her intent to "stay true" to the MMR course materials, "whatever they [ODE] want the lesson to look like." Fred and Gwen mentioned a similar intent, though they both acknowledged that ODE was not collecting data during the 2020–2021 academic year.

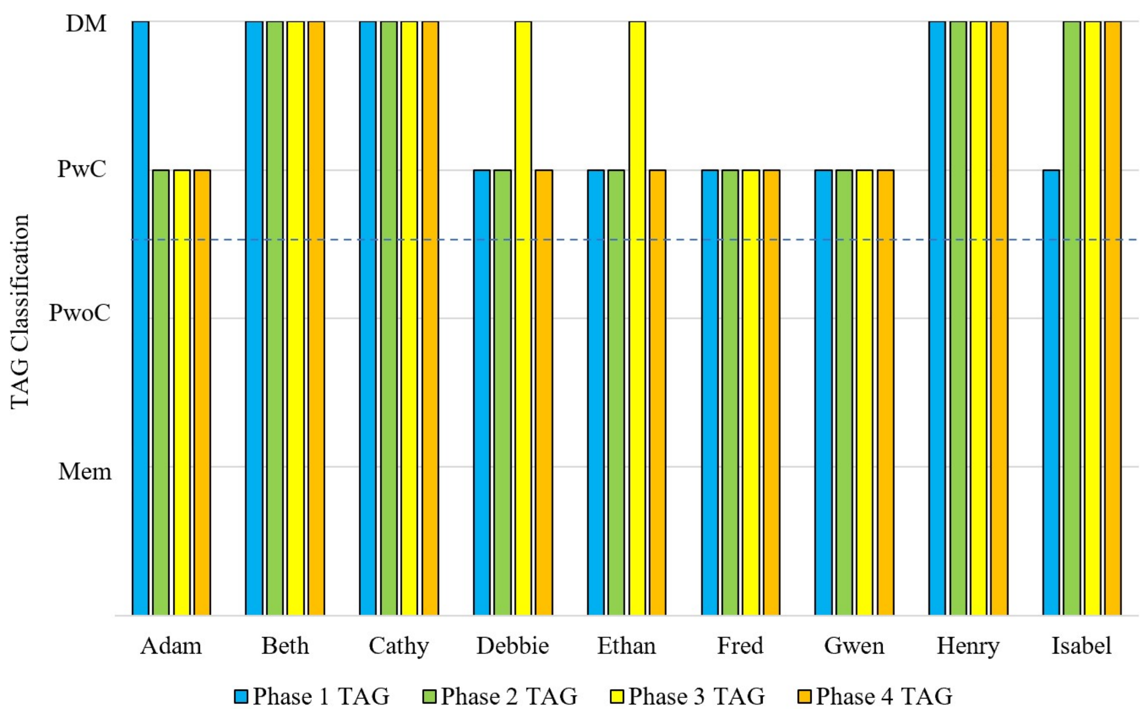
Trends in Teachers' TAG Classifications Across Task Phases

Teachers' self-reported TAG classifications changed less dynamically than the IQA scores identified by the researcher. As shown in Table 4 on page 163, 5 teachers provided the same TAG classification for their tasks across the four task phases: Cathy, Henry, and Beth categorized their tasks as *doing mathematics* for all four phases, whereas Fred and Gwen labeled theirs as *procedures with connections* consistently. Both Debbie and Ethan categorized their tasks as *procedures with connections* in Phases 1, 2, and 4, but felt that their tasks inclined to *doing mathematics* in Phase 3: Setup. Adam classified his task as *doing mathematics* at Phase 1: Selection but thought that it "increased" to *procedures with connections* from Phase 2: Planning onward. A visual representation summarizing this information is provided in Figure 7.

Like in Figure 6, the Phase 1: Selection–4 TAG classifications are represented as vertical bars, one for each task phase. The same four colors are used to signify each teacher's task at each of the four task phases: blue represents Phase 1: Selection, green represents Phase 2: Planning, yellow represents Phase 3: Setup, and orange represents Phase 4: Implementation. The vertical axis represents the four TAG classifications *memorization* (Mem), *procedures without connections* (PwoC), *procedures with connections* (PwC), and *doing mathematics* (DM), and each participating teacher is displayed along the horizontal axis. The horizontal dotted line segment indicates the separation between tasks that are considered high in cognitive demand (i.e., *procedures with connections* and *doing mathematics*) and those that are considered low in cognitive demand (*memorization* and *procedures without connections*). Using this representation, readers can track the progression of each teacher's task from Phase 1: Selection to Phase 4: Implementation.

Figure 7

Teachers' TAG Classifications Across the Phases of Their Tasks



Note. The dashed horizontal segment separates high cognitive demand tasks (TAG classifications of *procedures with connections* [PwC] and *doing mathematics* [DM]) from low cognitive demand tasks (TAG classifications of *memorization* [Mem] and *procedures without connections* [PwoC]).

In the following paragraphs, I discuss three trends in teachers' TAG classifications across the four task phases: (a) tasks that both inclined and declined in cognitive demand between phases, (b) tasks that strictly inclined in cognitive demand between phases, and (c) tasks that maintained cognitive demand between phases.

Tasks That Both Inclined and Declined in Cognitive Demand. As shown in Figure 6, Debbie and Ethan had both classified the "Follow the Bouncing Ball" task as a *procedures with connections* task in Phases 1–2. However, both teachers argued that the cognitive demand of the task increased to the level of *doing mathematics* in Phase 3: Setup, and then declined back to *procedures with connections* in Phase 4: Implementation. Debbie and Ethan identified similar changes in cognitive demand between Phases 2–3 and between Phases 3–4. Surprisingly, their reasoning for these changes was nearly identical as well. Recall Debbie's conclusion that "The more that you [the teacher] take out of it [a task], then the more the student has to come up with on their own, including the question. Then, you get more into the *doing mathematics*." She felt that the setup of the task inclined in cognitive demand because she allowed her students to investigate the relationship that they might explore and the potential variables that they could use and measure. However, guiding students toward the use of drop height and rebound height during implementation lowered the cognitive demand back to *procedures with connections* in Phase 4: Implementation.

Ethan's setup and implementation of the task followed a similar pattern, perhaps because both teachers followed the recommended instructional procedures in the MMR lesson plan. Ethan explained that Phase 3: Setup aligned more so with *doing mathematics* because his students led the discussion of what they noticed and what they wondered about the task. Moreover, he engaged his students in thinking, "What do I know, what do

I want to know, and how am I going to figure this out?" However, Ethan felt that the implementation of the task declined back into *procedures with connections* because his students simply followed the prompts on the task handout rather than exploring some of the ideas they generated on their own. In summary, both teachers thought that task setup aligned with *doing mathematics* because the task was more open-ended and student-led, whereas the implementation was at the level of *procedure with connections* because the students were directed toward a particular path instead of freely exploring their own ideas.

Tasks That Strictly Inclined in Cognitive Demand. Two teachers explained that their tasks inclined in cognitive demand in ways that changed their TAG classifications. Adam classified the "Remodeling Our Classroom" task at the level of *doing mathematics* for Phase 1: Selection, and his misunderstanding of the levels of cognitive demand influenced his reasoning that the task transitioned into the "higher" level of *procedures with connections* from Phase 2: Planning onward. On the other hand, Isabel provided sound reasoning for how the "Fitting a Line to Data - Earnings and Educational Attainment" (United States Census Bureau, 2021) task inclined from *procedures with connections* in Phase 1: Selection to *doing mathematics* in Phase 2: Planning.

Though both teachers conceptualized that their tasks increased in cognitive demand, this highlights a contrast based on their level of understanding of the TAG. Not only did Adam communicate a misunderstanding in the ranked order of *procedures with connections* and *doing mathematics*, but he also explained that "it [the cognitive demand of a task] always increases, because in the beginning, it [the task] is just a handout with instructions on it." He seems to understand cognitive demand as something that accumulates from one task phase to the next, increasing as students engage with the task: "All of these lessons grow from the initial standpoint… initially, they [the students] had to look over a handout and watch a video. That was it." Adam further explained that the cognitive demand increased as he and his students discussed the questions they had about the video, what they noticed, and what they wondered. Adam's analysis of the task shows a misconception about cognitive demand, as the cognitive demand of a task may decline from one phase to the next and does not accumulate from the first phase to the last.

Isabel's reasoning for her task's change in cognitive demand from Phase 1: Selection to Phase 2: Planning was grounded in evidence from the original task handout and the changes she made to it. Moreover, she referenced the TAG and pointed out specific criteria to justify the cognitive demand of the task at each phase. Though she felt that both versions of the task contained some elements of *procedures with connections* and *doing mathematics*, she classified Phase 1: Selection as the former and Phase 2: Planning as the latter. She acknowledged the focus on the procedures for graphing, writing equations, and creating lines of best fit in the original task, as well as the connections between tables, graphs, and equations. However, her adjustments to the task required students to access relevant knowledge and make appropriate use of it as they interpreted the real-world meaning of the *y*-intercepts and slope they found using linear regression. There was still a suggested pathway for students to follow, but the inclusion of various interpretation questions required students to analyze the task in ways that the original handout did not require. Though they both thought that their tasks inclined in cognitive demand, the cases of Adam and Isabel provide differing levels of understanding of the TAG and cognitive demand.

Tasks That Maintained Cognitive Demand. The remaining teachers classified each of their tasks at the same level of cognitive demand throughout the four task phases, as shown in Figure 6 on page 353. This suggests that these teachers did not identify any changes in the cognitive demand of their tasks from Phase 1: Selection to Phase 4: Implementation. However, Beth, Fred, and Gwen reported changes to each of their tasks that influenced their cognitive demand without justifying a change in TAG classifications.

Recall that Beth's students initially provided "superficial" written responses that she was not pleased with, so she provided feedback on the second day of instruction for them to revise and enhance the quality of their written work. Beth felt that her students' implementation fell into the *procedures with connections* realm somewhat during the first day, based on how their responses contained only numerical answers "without connections to what they were doing" outside of a particular problem. After providing feedback and listening to the group discussions that occurred during the second day, Beth thought that her students' implementation of the task rose back into the *doing mathematics* area and met the potential of the task in Phases 1–3. Though she classified the task as *doing mathematics* in all four phases, she recognized changes in cognitive demand within a single phase, Phase 4: Implementation.

Fred and Gwen's cases are different than Beth's, but more relatable to each other. Both teachers addressed changes to their respective tasks in Phase 2: Planning that increased the cognitive demand, but not quite enough to elevate their tasks from *procedures with connections* to *doing mathematics.* Fred revised his modified version of the task again after our first interview, adding questions that encouraged students to interpret the slope and *y*-intercept of their linear models. Adding these questions, according to Fred, would "drive up" the cognitive demand of the task; however, the task still remained "on the edge" of *doing mathemati*cs and in the *procedures with connections* level because the task connects a context to students' prior knowledge of slope and linear equations. Instead of adding to a task, Gwen made a subtraction from the "Discovering Slope" task; though it was not the task I observed her teach, Gwen described this change to the task as part of her planning for the Ramps Context. She had considered including the first page of the "Discovering Slope" handout but decided against it because "it implies that there will be direct instruction." Gwen wanted her students to ask questions rather than reviewing the concept of slope prior to their engagement in the "Ramp It Up" (King, 2015) task; removing this first page increased the cognitive demand of the task according to Gwen, but not enough to justify a change in TAG classification.

Mismatches: Comparisons Between IQA Scores and TAG Classifications

A mismatch in task analysis occurred when the researchers and a teacher reach opposing conclusions for whether the cognitive demand of a task is high at one of the four task phases. High cognitive demand tasks, according to the IQA rubrics, are those that score a 3 or 4. Similarly, tasks that are classified as *procedures with connections* and *doing mathematics* using the TAG are also considered to have high cognitive demand. Otherwise, a task is considered to have low cognitive demand. As shown in Tables 5, 7, and 11 on pages 165, 202, and 267, three mismatches occurred: Adam's task, Cathy's task, and Gwen's task at Phase 4: Implementation scored 2 on the IQA, whereas these teachers classified their tasks as *procedures with connections*, *doing mathematics*, and *procedures with connections*, respectively. This information is summarized in Table 36.

Table 36

Mismatches in Task Analysis

Teacher	Score Level	IQA Score	TAG Classification
Adam	Phase 4	2	PwC
Cathy	Phase 4	2	DM
Gwen	Phase 4	2	PwC

The three mismatches occurred in Phase 4: Implementation, indicating a difference in the analysis of how each task unfolded with students. As I described in the analysis of each teacher's task implementation, I assigned scores of 2 on the IQA primarily because students' engagement focused on procedural aspects of each task. Adam's students calculated the surface area of their classroom, made necessary unit conversions to determine the amount of paint they would need to cover the walls, and determined the cost of paint and other materials they would need for the project. However, they divided the work of the task among themselves (with Adam's instruction to do so) and rarely communicated any evidence of conceptual reasoning and understanding. Cathy's students collected data, ran a linear regression to determine a line of best-fit, and used their models to predict the number of candies that could be grabbed given the size of one's hand. But it was Cathy who voiced some of the conclusions that were to be drawn based on their findings, rather than the students themselves. Finally, Gwen's students routinized the work of their task by removing the real-world context and using the procedures they had learned in Chemistry to set up and solve proportional equations. The focus of the task shifted from problem-solving and making connections to simply finding answers.

Of the three teachers, Cathy and Gwen seemed to notice some decline in the cognitive demand of task implementation; however, they did not adjust their TAG classifications to reflect what they had observed. Of the two, Gwen was even more aware, explaining that her students had learned how solve problems involving similar triangles in Chemistry beforehand. Through her students' conversations, she identified that their Chemistry teacher had "taught them proportions; taught them to cross multiply and divide." Gwen felt that the procedure was "ingrained in them [the students] from hearing it over and over." She explained her attempt at getting students to visualize what was given in each problem and how to think through it, but they leaned on what they had recently done in Chemistry instead. Gwen acknowledged that her students likely did not understand the procedures they used and why they were using them. "They [the students] were kind of alluding to the connection between similar triangles and proportions… but they didn't really have the connection," she concluded.

Cathy similarly acknowledged some level of decline in cognitive demand during task implementation. For example, she thought that her students needed to improve at self-monitoring and self-regulating their own cognitive processes because some students left parts of their handout blank without asking questions and seeking help. Cathy also explained that "Based on their [her students'] responses, I don't think they stretched themselves as much as I would have liked," communicating an awareness that her students could have done more thinking and reasoning with the task. As she considered whether her students had been required to put forth considerable cognitive effort, Cathy concluded, "I'm not sure that this [the task] stretched them much.

Moreover, Cathy said that her students "could have done more" to explore and understand the nature of mathematical concepts, processes, and relationships. She felt that they did not experience much productive struggle throughout their engagement with the task because they simply took measurements, collected data, and generated a model using technology. Cathy attributed this partially to her "benefitting from a good group of kids, and they didn't need to struggle." Indeed, it appeared as though her students did not struggle with the task, but it was because Cathy did the analysis and interpretation of students' linear models. For example, she stated that the combined class model was "not a good fit" based on students' data. The students did not discuss the mathematical model and only used it to "plug in" values that the teacher told them to make predictions. Though less specific than Gwen, Cathy similarly attributed students' thinking to their previous mathematics learning experiences: “They [the students] have just been through 12 years of math education that is rather procedural,” encouraging students to think, “Which thing that someone told me to do before, should I do now?” rather than thinking about problems conceptually.

Despite acknowledging a potential decline in cognitive demand, Gwen and Cathy maintained the ratings they had assigned for their tasks in Phase 3: Setup. Neither teacher lowered their classification of the task to *procedures without connections*, resulting in a mismatch in task analysis. Gwen maintained her Phase 3: Setup rating of *procedures with connections* because she felt that her attempts to “push” her students in that direction were enough to maintain the cognitive demand of the task. “Even though it may not have been through work of their own, I was pushing them to get to that point,” she explained, justifying her TAG classification for Phase 4: Implementation. Gwen explained that her students were “alluding to the connection between similar triangles and proportions,” but then recanted, saying, “but they didn’t really have the connection… so I guess that doesn’t quite count for that specific scenario.” Though Cathy’s final classification for Phase 4: Implementation was *doing mathematics*, she felt that her students were “somewhere between the two” levels of *doing mathematics* and *procedures with connections* and “jumped back and forth” among them. At one point, she stated that “they [the students] might still be more at the *procedures with connections* level as far as their thinking.” However, she concluded, “I think that I kept more of the bullets than I lost,” regarding the criteria listed for *doing mathematics* in the TAG.

Unlike Cathy and Gwen, Adam did not express any potential decline in cognitive demand with the implementation of his task. On the contrary, he only expressed positivity in response to his students' work during the two lessons. To Adam, his students were making connections by "finding different ways to represent data" and recognizing that some aspects of the task were not realistic. Adam acknowledged that the student-centered nature of the task and "allowing them to drive the vehicle" contributed to its high cognitive demand; however, the decisions that students made regarding the task were nonmathematical, such as what objects to include in their remodeled classroom. The decisions that they made did not concern various problem-solving strategies or mathematical processes to use. Adam also attributed the high cognitive demand of the task to students' collaboration with each other:

> They [students] are bringing others in their group into the task to say, "Okay, now help me deal with this. Get a deeper understanding of what I'm trying to do to make sure that it's correct." … Again, it's that sense of community, and they're making that connection.

This might occur frequently in Adam's class, but was not something that I observed as the students worked on their own sections of the task. On one or two occasions, I heard one student ask another to check their work, but this did not occur frequently; the students were mostly silent at work individually. Moreover, making connections to each other's work in this way does not contribute to making *mathematical* connections between concepts and representations, as is intended in the TAG.

Among the mismatches in task analysis, an interesting pattern is that 2 of the 3 teachers had not taught the MMR course before. Of the 8 MMR teachers involved in the present study, only Adam and Cathy were first-year pilot teachers. Gwen, on the other hand, was a second-year pilot teacher and had taught MMR before. Though this pattern might be a coincidence, it suggests that the teachers most unfamiliar with the MMR course and the associated PD might have had more difficulty maintaining the high cognitive demand of their tasks. However, it also suggests that teachers with more experience teaching MMR might be more likely to maintain the cognitive demand of their tasks; of the 6 second- and third-year MMR teachers, only Gwen's task declined in cognitive demand during task implementation. Multiple experiences with the MMR course and PD might have influenced this trend, though more research is needed to draw such a conclusion. With the vast amount of material that the first-year MMR teachers were exposed to (i.e., new curricular materials, mathematics content, and pedagogical approaches), in addition to challenges teachers faced during the COVID-19 pandemic, it is unsurprising that their tasks declined in cognitive demand more so than second- and third-year MMR teachers' tasks.

Chapter Summary

I discussed the research findings in this chapter, including conclusions derived from the analysis of interviews, observations, and mathematical tasks. I presented the 9 individual cases for each participating teacher, including (a) the analysis of each teacher's task across the four task phases using the IQA, (b) teachers' analysis of the task at each phase using the TAG, and (c) teachers' reasons for selecting, planning, setting up, and implementing tasks. The findings across teachers showed that, though they each provided their own views of task selection, planning, setup, and implementation, there were some commonalities especially among the 8 MMR teachers. They selected MMR tasks for use in their classrooms and attended to the lesson plans provided by ODE and typically included instructional elements and practices emphasized in the PD for the course (e.g., brainstorming sessions with students, collaborative group work, and student discussions). In response to changes in the instructional environment due to the COVID-19 pandemic, the teachers modified tasks to accommodate those changes. The cognitive demand of tasks tended to stay high, except in three cases where the teachers provided too much guidance and the focus shifted to executing procedures to find answers.

Chapter 5: Discussion, Recommendations, and Conclusion

This chapter contains four main sections plus a summary and conclusion. In the first main section, I restate the research questions and discuss how the data address these questions. The next section includes theoretical and practical implications for the fields of mathematics education and teacher education. In the third and fourth sections, I provide recommendations for future practice and future research.

Answers to the Research Questions

Three research questions guided the present study. The first question focused on how high school mathematics teachers selected, planned, set up, and implemented instructional tasks and teachers' reasons for their actions across these four phases. The second question addressed teachers' perspectives on how the cognitive demand of their tasks changed across the four task phases. The third question pertained to teachers' analysis of cognitive demand of a task and the mismatches between the researcher and co-observer's analysis and the teacher's analysis of each task. The three research questions follow:

1. When attempting to use a high cognitive demand mathematical task, what actions do high school mathematics teachers take, and for what reasons, while
    a. selecting the task from written source materials?
    b. planning the task for use with their students?
    c. setting up the task immediately prior to student engagement?
    d. implementing the task as students engage with it?

2. What reasons do high school mathematics teachers give to explain the change in the cognitive demand of a task across the four phases of selecting, planning, setting up, and implementing the task?
3. What reasons do high school mathematics teachers give for assessing the cognitive demand of a task at each phase, and in particular, what reasons do they give when there is a mismatch between a teacher's assessment of the cognitive demand of a task and the researchers' assessment of that phase of the same task?

In the following paragraphs, I explain the answers to these three research questions especially concerning the themes and trends identified in Chapter 4.

Selecting Tasks

The 18 themes that were presented and discussed in Chapter 4 answer the first research question. The factors teachers cited as affecting their selection of tasks were (a) considering face-to-face, remote, and hybrid learning formats; (b) promoting active student engagement; (c) addressing important mathematics content, skills, and processes; (d) connecting to real-world contexts; and (e) considering past success.

Unsurprisingly, the participating teachers chose tasks that were suitable or adaptable for remote instruction when given the freedom to choose which MMR tasks they would use. Those that involved extensive, hands-on data collection procedures were avoided by some teachers because their students could not experience the task to the full potential as they might in a typical face-to-face setting. The tasks that were chosen promoted active student engagement through explorations and opportunities for student discussion, aspects of the tasks that the teachers highlighted. Tasks such as "Follow the

Bouncing Ball" and "Starbursts™ Grab" were used to develop students' understanding of linear functions, but more importantly, the participating teachers emphasized that such tasks encouraged students to problem-solve, think critically, and analyze mathematical relationships. Additionally, the tasks were selected to help students connect mathematics to their daily lives. The MMR teachers' selection of tasks appears to be aligned with the purpose and goals of the course and associated PD; however, the teachers also relied on the use of tasks that had been successful in the past, confident that they would be successful again.

Planning Tasks

The factors teachers cited as affecting their planning for using tasks with students were (a) flexibility; (b) accomplishing goals and objectives; (c) desire to maintain fidelity to the provided MMR lesson plans; (d) anticipating challenges and student responses; and (e) adjusting tasks for face-to-face, remote, and hybrid learning formats.

The participating teachers were flexible in their planning, understanding that the length or duration of a task is determined by the students; this is especially the case when tasks are open ended, and the students have some autonomy in determining what aspects to explore. Tasks were specifically modified so that students would achieve certain learning goals, such as proficiency with spreadsheets or how to interpret coefficients in a linear model. As might be expected, the MMR teachers expressed a desire to follow the lesson plans that were provided in their PD program closely. Those teaching MMR for the first time especially maintained the desired flow of the lessons and made sure to incorporate the recommended instructional procedures and materials. Some teachers, like

Fred, were more willing to adapt due to their familiarity with the tasks, though also acknowledging that ODE desired consistency when conducting research on the MMR course.

The tasks challenged teachers to anticipate difficulties and student responses of various types. Teachers anticipated students' engagement with a task, especially in remote and hybrid learning environments because these were unfamiliar territory for the teachers. Even those teaching face-to-face were compelled to consider how they might set up and implement tasks safely, minimizing health and safety risks during the COVID-19 pandemic. Furthermore, the open-ended nature of the tasks required teachers to consider how students might respond and how they might follow-up on students' responses. The more student-centered a task is, the more opportunities for it to diverge from a single, predetermined pathway; because many of the tasks encouraged students to explore and consider problem-solving situations, the MMR teachers had to anticipate how their students would interact. The teachers also considered how they might support their students in understanding the mathematics; several teachers explained that their MMR students had typically experienced limited success with mathematics in the past.

Setting Up Tasks

The factors teachers cited as affecting their task setup were (a) asking, "What do you notice?" and "What do you wonder?"; (b) facilitating whole-class discussions before student work time; and (c) teachers communicating their expectations.

The inclusion of a brainstorming stage and a preliminary discussion are features of MMR tasks that are typically included in the lesson plans developed by ODE. In the

PD for the course, the MMR teachers were encouraged to adopt these instructional strategies as ways for their students to engage in mathematical thinking and problem-solving. The three tasks that did not include preliminary discussions, namely Beth's, Henry's, and Isabel's, differed from the others because they did not require students to investigate a real-world situation and collect data. The remaining 5 teachers sought to access students' relevant prior knowledge about the real-world situations and the mathematics of their tasks as a starting point for task implementation. Many teachers aimed to set a standard for students' engagement with the tasks and the work that their students produced. By communicating their expectations at the beginning, teachers hoped that their students would engage more deeply and produce higher quality work than they might have otherwise.

Implementing Tasks

The factors teachers cited as affecting their task implementation were (a) encouraging productive struggle, (b) eliciting evidence of students' thinking and reasoning, (c) monitoring and facilitating student engagement, (d) asking questions, and (e) providing instructional support.

Productive struggle is a concept that was introduced to teachers during the MMR PD and was one of many pedagogical approaches that was endorsed; not all the participating MMR teachers used the specific term "productive struggle," but many of them spoke positively about challenging their students to think critically, make mistakes, and learn from them. I cannot claim that the MMR PD directly influenced these teachers to feel this way, but this provides evidence that the PD might have had an impact on teachers' beliefs about productive struggle. Similarly, the MMR teachers emphasized a desire for their students to communicate their thinking and reasoning; this was another focus of the PD. Other factors may be responsible for teachers' attitudes, but this is still a promising outcome.

The participating teachers highlighted three instructional processes, leading to the development of themes for facilitating students' engagement, asking questions, and providing support. When facilitating students' engagement with the tasks, teachers actively listened to student-to-student conversations for evidence that they might be struggling with a particular concept or idea. Doing so is essential with high cognitive demand tasks because the cognitive demand may decline if students fail to make mathematical connections and develop conceptual understanding (Stein et al., 2009). This was the case with Adam, who rarely checked in with his students except to ask them what they were working on and verify that they were making progress. Adam missed opportunities to ask his students conceptual questions about the procedures they were using by remaining silent. This shows that there is a fine line between saying too much and saying too little; the former might limit students' opportunity to think, whereas the latter might result in missed opportunities to expand on students' thinking.

The teachers frequently asked questions to scaffold and assess students' thinking, maintaining the emphasis on student thinking rather than taking over the thinking for themselves. Effective teacher questioning was indeed one of the instructional practices that helped to maintain high cognitive demand for many teachers' tasks. For Henry's task, it even resulted in an increase in cognitive demand because his questions elicited evidence of students' thinking and reasoning beyond the scope of the original task. In addition, the participating teachers identified reasons for providing instructional support and typically did so in ways that did not lower the cognitive demand of their tasks. Their assistance frequently allowed students to visualize or conceptualize ideas in another way without taking over their thinking. Even Gwen, whose task declined in cognitive demand because the focus shifted to correct answers, instructed her students to create visual diagrams to help them picture the differences between ramps with various slopes. The cognitive demand of the task declined, but for other reasons.

Change in Cognitive Demand Between Task Phases

In general, the participating teachers provided reasonable justification for their analysis of the cognitive demand of their tasks in Phases 1–2, Task Selection and Task Planning. They understood, from the onset, that the tasks they used were in the higher-level categories of *procedures with connections* and *doing mathematics*. They referred to the TAG when prompted to do so and cited specific parts of a task that aligned with the given criteria in the guide. Debbie's understanding of *doing mathematics*, for example, was exceptional; she acknowledged that students do mathematics when they determine what questions to ask, how they will answer their questions, and what variables are relevant to explore. Fred pointed out that tasks with real-world contexts are not necessarily high in cognitive demand, identifying a "superficial feature" (Stein et al., 2009, p. 7) that some teachers erroneously associate with high cognitive demand.

Fred and Isabel acknowledged a common increase in cognitive demand from Phase 1: Selection to Phase 2: Planning by adding prompts for their students to interpret the real-world meaning of the coefficients in their linear models. However, Fred did not feel that this change was enough to increase the cognitive demand from *procedures with connections* to *doing mathematics,* but Isabel did. Gwen categorized the “Ramp It Up” task as *procedures with connections* at the first two phases, and she recognized that removing the first page of the “Discovering Slope” handout would increase the cognitive demand of the task because the first page provided students with solution pathways and algorithms upfront.

Adam, on the other hand, differed from the other teachers significantly; it appears that he perceives *procedures with connections* at a level above *doing mathematics*, as he said his task would be “higher” in cognitive demand at the latter stage. Moreover, he stated that cognitive demand “always increases,” which is false because cognitive demand may decrease or be maintained between any two task phases. Adam’s perspective was unique in this study but may be characteristic of other teachers who misinterpret the TAG and the levels of cognitive demand. The remaining teachers made few changes to their tasks between Phases 1 and 2, Task Selection and Task Planning, aside from reformatting their tasks to accommodate remote instruction when they felt it necessary.

Similarly, few teachers reported changes in cognitive demand between task Phases 2 and 3, Task Planning and Task Setup. Adam, Beth, Cathy, Fred, Gwen, Henry, and Isabel classified their tasks at the same level of cognitive demand in Phase 3: Setup as they had in Phase 2: Planning, though not all these teachers’ tasks truly maintained the same level of rigor. Indeed, Beth, Henry, and Isabel’s tasks maintained cognitive demand during task setup because the setup phase consisted only of providing materials to students and setting them free to work. Despite the fact that Cathy and Fred initiated whole-class discussions prior to student work time, the cognitive demand of each of their tasks did not increase.

Adam and Gwen's tasks changed in ways that they did not acknowledge using the TAG; Adam's task was not scored for Phase 3: Setup because it was not observed, but he described how he walked his students through the process of determining the length of the piping that would go inside his remodeled classroom based on the task handout. This instruction likely lowered the cognitive demand of the task, as students no longer needed to determine the spatial positioning of the pipeline on their own. Alternatively, Gwen prompted her students to "verbalize" their approach to solving problems when she passed out the task handout, a move that enhanced the cognitive demand of the task because explanations were not explicitly called for on the handout. This shows that there were instances where teachers failed to either notice or acknowledge changes in cognitive demand between Phases 2 and 3, Task Planning and Task Setup. This is perhaps because classifying the cognitive demand of a task during instruction with students in Phases 3 and 4, Task Setup and Task Implementation, is more challenging than identifying the potential of a task in Phases 1 and 2. Tasks become more complex as they inspire interactions between teachers and students, making it more difficult to determine their level of cognitive demand using the TAG accurately.

Only 2 participants addressed changes in cognitive demand between Phases 2–3, Task Planning and Task Setup: Debbie and Ethan. Both teachers used the same task and identified a similar reason for change in cognitive demand: Each stated that the cognitive demand increased from *procedures with connections* in task planning to *doing mathematics* in task setup because the discussions in the setup phase allowed their students to determine what variables to explore, how they might measure them, and how they might use this information to complete the task. However, these are features of the "Follow the Bouncing Ball" task that were also evident in Phases 1 and 2, Task Selection and Task Planning. In fact, the task handout itself (Phase 1) prompted students to do what Debbie and Ethan had them do during task setup (Phase 3). Therefore, the cognitive demand was already at the level of *doing mathematics* in Phases 1–2, and the cognitive demand did not change during task setup. The cognitive demand might have changed through the specific interactions between each teacher and their students, but simply having them brainstorm variables of interest did not introduce a change to the cognitive demand of the task. So, not only did some teachers fail to acknowledge changes in cognitive demand from Phase 2: Planning to Phase 3: Setup, but some even saw changes when they did not necessarily exist.

Concerning the transition between Phase 3: Setup to Phase 4: Implementation, both Debbie and Ethan reported that their tasks were implemented at the level of *procedures with connections*. Debbie's reasoning focused on the connections that her students made between the various mathematical representations of graphs, equations, and numerical data, though she did not specifically comment on how her students' implementation of the task failed to satisfy the requirements for *doing mathematics*. Ethan, on the other hand, felt that the task had the potential to reach the level of *doing* *mathematics* during implementation, but felt that his students were not yet at that level. His reasoning was that he did not have “the time and the ability” to guide students toward what the MMR lesson plans called for, especially because his students were not able to do some of the other linear modeling tasks such as “Starbursts™ Grab” and “Bungee Barbie.” Ethan implied that the restrictions with remote learning limited the capability for his task to reach its full potential.

Though Beth did not classify her task at different levels of cognitive demand between Phases 3–4, Task Setup and Task Implementation, she noted that the cognitive demand of her task fluctuated within Phase 4: her students’ written responses were “superficial” in her eyes and lacked detailed reasoning after the first day, so she provided oral feedback at the start of the second day asking them to revise and explain their work in more detail. After the second day, Beth felt that her students had gone back and revised their work to her satisfaction, resulting in the classification of *doing mathematics* for Phase 4 overall.

Among the participants, 6 teachers reported identical TAG classifications from the former task phase to the latter. Of these teachers, Adam, Cathy, and Gwen’s tasks declined in cognitive demand and are discussed in the following part of this section. Fred’s task maintained its cognitive demand according to the IQA rubrics, whereas Isabel’s task was only implemented by two students at their own pace and could not be analyzed using the IQA. Henry did not acknowledge the increase in the cognitive demand of his task from Phase 3 to Phase 4; he classified the task as *doing mathematics* across all four phases, though his students were pressed for and provided explanations that were not explicitly called for by the task. Henry might not have noticed this difference between task phases, but it is also worth mentioning that the TAG is not as explicit as the IQA at addressing evidence of students' thinking and reasoning. The IQA clearly cites this as a distinction between a score of 3 and a score of 4, whereas the TAG uses different criteria entirely, possibly explaining why my co-observer and I noticed this increase in cognitive demand.

Mismatches in Task Analysis

A mismatch in task analysis occurred when either the researchers or a teacher identified a task as high in cognitive demand (i.e., scores of 3 and 4 in the IQA or *procedures with connections* and *doing mathematics* in the TAG) and the other identified a task as low in cognitive demand (i.e., scores of 0, 1, and 2 in the IQA or *memorization* and *procedures without connections* in the TAG). Three mismatches in task analysis occurred, each during task implementation: (a) Adam classified his task as *doing mathematics*, (b) Cathy classified her task as *doing mathematics*, and (c) Gwen classified her task as *procedures with connections*, whereas all three tasks scored a 2 on the IQA.

As discussed in the previous paragraphs, the participating teachers acknowledged fewer changes in cognitive demand among Phases 3 and 4, Task Setup and Task Implementation, than they did among Phases 1 and 2, Task Selection and Task Planning. This could be because the third and fourth task phases are greater in complexity than the first and second. Therefore, it is not entirely surprising that the only mismatches occurred in the implementation phase. This could also be explained by the fact that the teachers made few substantial changes to their tasks when planning and the teachers who included a distinct setup phase were guided by the MMR handouts and lesson plans provided by ODE.

Of the 3 teachers whose tasks scored a 2 on the IQA for Phase 4: Implementation, Adam's reasoning for classifying his task as high in cognitive demand differed from Cathy and Gwen's. One distinguishing characteristic of Adam's reasoning was that he misunderstood the TAG levels of cognitive demand—essentially reversing the ranking of *procedures with connections* and *doing mathematics*. Moreover, his reasoning for why the implementation of the task was at the *procedures with connections* level addressed real-world connections that his students made rather than the mathematical connections they made. One such example was students' realization that some aspects of the task were not practical in real-life, like the inclusion of a fireplace and carpet in a school classroom. Adam also expressed with enthusiasm that his students were building a sense of community and collaboration with various aspects of the task, helping and supporting each other when they struggled and keeping each other on task. Though these are positive outcomes when engaging students with a task, they do not influence the cognitive demand of a task unless they address mathematical understanding and sensemaking; based on these data, Adam's conception of cognitive demand appears to be still developing, especially concerning task implementation.

Unlike Adam, both Cathy and Gwen acknowledged some degree of decline in the cognitive demand of their tasks between Phases 3 and 4, Task Setup and Task Implementation. However, neither lowered the TAG classifications from what they provided in Phase 3 to Phase 4. Gwen explained that her students did not verbalize their thinking and did not make connections to the level of her expectations, relying on the procedures they had learned in Chemistry for working with similar triangles without understanding how and why they made sense mathematically. She concluded that her students were at the level of *procedures with connections*, "though it may not have been through work of their own" as she had to "push them that point." The reason for the mismatch, in Gwen's case, seems to be that she classified the task based on the level of thinking that she *attempted* to have her students reach, not necessarily their *actual* level of understanding. Gwen seemed to have a stronger understanding of how her students engaged with the task than Cathy, though she did not change her TAG classification to match it. She acknowledged gaps in her students' understanding but might also have been unwilling to admit that the task had declined into the low cognitive demand space.

Cathy might also have been unwilling to acknowledge a decline in cognitive demand. This is because, like Gwen, she communicated dissatisfaction with her students' engagement with the task. Cathy expressed that her students could have "done more" to understand the nature of mathematical concepts, processes, and relationships, and that there was "not much struggle left" during implementation. Though she referenced specific criteria for *doing mathematics* in the TAG that she felt were lacking, she concluded that she "kept more of the bullets than she lost," resulting in the same level of cognitive demand as in Phase 3: Setup. The reason for the mismatch with Cathy appears to stem from this apparent contradiction, though I cannot make a claim as to the reason for it. Perhaps she felt that her students' implementation of the task satisfied the TAG criteria for *doing mathematics* enough to warrant the classification, despite lacking in some respects.

Teachers' experience with the MMR course and the associated PD might have contributed to the mismatches in task analysis that occurred. As I mentioned in Chapter 4, Adam and Cathy taught the MMR course for the first time during the 2020–2021 academic year, whereas Gwen had taught the course once before during the 2019–2020 school year. Of these 3 teachers with mismatches, Adam and Cathy were less familiar with the curricular materials, mathematics content, and pedagogical approaches of the MMR course than the other 5 MMR teachers. Moreover, they might have also been less familiar with the MTF and the TAG than teachers who had experienced the MMR PD prior to the 2020–2021 academic year. Therefore, it is not surprising that Adam and Cathy's tasks declined in cognitive demand and their analyses did not match with the researcher and co-observer's. The case of Gwen provides evidence that more experienced MMR teachers might still have trouble maintaining the cognitive demand of their tasks and disagree with researchers' task analysis. However, it seems that more experienced MMR teachers are more likely to maintain high cognitive demand and analyze tasks more accurately.

Theoretical and Practical Implications

This section presents a discussion of how the findings of this dissertation research have the potential to influence the fields of mathematics education and teacher education. First, I discuss the Stein and colleagues (2009) MTF and the expanded four-phase model of the phases in the life of a task that I developed to guide the present study. Then, I explain how teachers' task use in the present study relates to the findings of previous empirical studies on teachers' use of tasks. Next, I address teachers' focus on active student engagement, followed by a section concerning the relationship between the IQA and TAG—the two rubrics that were used to analyze the cognitive demand of instructional tasks.

The Phases in the Life of a Mathematical Task

The QUASAR research of the 1990s led to the development of the MTF, a framework that describes how tasks progress from their appearance in instructional materials to how they are set up by teachers and enacted by students (Stein et al., 1996, 2009). Each phase of selecting, setting up, and implementing a task is influential in determining what students learn from the task. I hypothesized that a *planning* phase might also be influential because teachers do not always set up and implement tasks as they appear in curricular materials; they sometimes add to, remove from, or revise aspects of a task as they prepare for teaching the task (Earnest & Amador, 2019; Grouws et al., 2013). Although Smith and colleagues (2008) claimed that teacher planning affects how teachers set up tasks prior to student engagement, there had been no empirical evidence to date to support this claim. This led to the development of the four-phase task progression depicted in Figure 1 on page 50.

Of the 9 participating teachers, only Isabel chose to adjust her original task in ways that changed its IQA score for Task Potential. In particular, Isabel modified the task to change it from a score of 3 for Phase 1: Selection to a score of 4 for Phase 2: Planning. Although the remaining 8 teachers, all of whom were teaching the MMR course, adapted tasks to make them suitable for remote instruction or safe to implement in face-to-face settings during the COVID-19 pandemic, their tasks retained identical IQA scores between the selection and planning phases. It is possible, perhaps even likely, that the 8 MMR teachers felt compelled to remain true to the tasks as written in the state of Ohio instructional materials because they were participating in a statewide pilot of these materials. Under other circumstances, they might have made greater adjustments during the planning phase that would have affected the rating of the cognitive demand of the task. The IQA scores for teachers' tasks in selection and planning might have changed more dramatically if I had recruited more teachers outside the MMR program. Nevertheless, the findings of this study suggest that the planning phase is influential in the progression of a task because of (a) the many ways that teachers adjusted their tasks and planned around the COVID-19 pandemic and (b) the careful consideration that teachers took when anticipating task setup and implementation.

An unexpected outcome of the present study was the degree of attention that teachers made to adjust to the pandemic when planning for instruction. This unprecedented situation presented teachers with numerous challenges that they endeavored to overcome, such as finding ways to engage students in collecting real-world data in remote learning environments and collecting data safely in face-to-face settings. The context of the MMR course highlighted teachers' efforts to plan for meaningful mathematical work and discussions in remote settings, whereas the teachers using lecture or other teacher-centered pedagogies might have transitioned to COVID with fewer adjustments to their tasks. Providing online lectures and lecture notes, for example, is simpler than engaging students actively in online learning settings.

The inclusion of a planning phase in the four-phase model also illuminated aspects of teachers' task use that were not identified in the other three phases: for example, many of the participating teachers explained their anticipation of students' responses and challenges when planning their instruction. To prepare for how they might set up and implement the task with students, teachers such as Henry considered what questions their students might ask and ways that they might respond to support students' engagement and understanding. When planning her instruction, Gwen highlighted mathematical concepts and ideas that she felt that her students would struggle with as they engaged with the task. Anticipating these potential student difficulties might have influenced her interactions with students when they struggled in class.

The findings of the present study also highlight the importance of teacher reflection after attempting to use a high cognitive demand task. Through pre- and post-observation interviews, participating teachers reflected on their task use in ways that they might not have otherwise been able to. Many teachers had only several minutes to debrief between one class period and the next, and the practical work of teaching typically provides limited time for teachers to deeply consider their past instruction. Even in the pre- and post-observation interviews, participating teachers took a practical approach to reflecting on their practice. They provided evidence of their reasoning when asked about task selection, but discussions about planning, setting up, and implementing tasks focused primarily on teachers' *actions* rather than the underlying *reasons* for such actions.

Post-observation interviews, in particular, which involved task setup and implementation, centered around teachers' experiences and decision-making in the moment rather than connections to theory (i.e., the MTF & the TAG). Student learning is the final stage in the Stein et al. (2009) MTF, but I suggest also that *teacher reflection* is an essential component that should follow the implementation of a task. To move the field of mathematics teacher education forward in the use of high cognitive demand tasks, it is crucial that teachers reflect on their task use from both practical and theoretical lenses. For teachers to do so effectively, they must be provided with the time and resources to reflect deeply on their instruction in meaningful ways.

Maintaining the Cognitive Demand of a Task

The QUASAR studies also highlighted that the cognitive demand of a task tends to decline from its appearance in source materials to implementation with students (Stein et al., 1996; Henningsen & Stein, 1997). Since this seminal QUASAR work, numerous studies have had similar findings (e.g., Boston & Smith, 2009; Jackson et al., 2013; Ni et al., 2018). In the present study, 4 of the MMR teachers employed tasks in such a way that a decline occurred in the cognitive demand from Phase 1: Selection to Phase 4: Implementation; whereas, 3 teachers maintained the cognitive level of their tasks, and 1 inclined the level of the task.

It is encouraging that half of the MMR teachers were able to maintain or increase the cognitive demand of the tasks across their lifespans. Nonetheless, like past research, a decline in cognitive demand was also common. Stein and colleagues (1996) found not only that a decline in cognitive demand was common but even that *challenges became* *nonproblems*, in other words, that the teachers dramatically reduced the amount of work and thinking required of students across the lifespan of a task. By contrast, even though many of the MMR teachers described their desire for students to engage in productive struggle, some had a drop in cognitive demand but not in such a dramatic fashion as found by Stein and colleagues. For example, a cognitive-level drop occurred when Adam supplied students with the placement of the gas furnace in their remodeled classroom and provided example scale drawings for students to refer to.

Focus shifting to correct answers was another of the common reasons for cognitive demand to decline in previous studies (Henningsen & Stein, 1997). This occurred for Cathy, who directed students' attention to "plugging" numerical values into their linear models and "seeing what they get" without discussing the meaning behind what they were doing and interpreting their results. In Gwen's case, her students shifted the focus of the task toward using procedures they had learned in Chemistry without understanding how and why they were using them.

Among the factors associated with a maintenance in cognitive demand identified by Henningsen and Stein (1997), teachers in the present study (a) selected tasks that built on students' prior knowledge, (b) provided scaffolding, (c) pressed students for explanations and meaning, and (d) drew conceptual connections. These practices occurred throughout the setup and implementation of tasks that maintained cognitive demand. For instance, the tasks focusing on linear equations built on what students had learned about linear functions both in previous MMR tasks and in Algebra 1. Participating teachers also provided instructional support without doing the work and thinking for students by prompting them to create visual diagrams. Many of the teachers frequently questioned their students and pressed for explanations, even to the point of enhancing the cognitive demand of the task in the case of Henry. Fred and Isabel, among others, drew conceptual connections by prompting students to interpret the meaning of their linear models and the numerical coefficients within them.

Though the teachers involved in the present study typically reinforced the findings of Henningsen and Stein (1997), it is worth mentioning that they also introduced factors influencing a change in the nature of high-level tasks. Namely, Henry and Gwen prompted students to explain their thinking and reasoning verbally when their written tasks did not do so, resulting in increases in IQA scores from 3 to 4 in both instances, though Gwen's increase was in the setup phase and Henry's was in implementation. Gwen set the expectation for her students to "verbalize their approach" to solving each problem, but her students failed to do so during implementation and the cognitive demand of the task declined as a result.

A promising finding is that 5 of 8 MMR teachers involved in the present study implemented their tasks at high cognitive levels, whereas only three tasks experienced a decline into *procedures without connections*. This is encouraging because of how difficult it can be to maintain the cognitive demand of a task in practice. In addition, all the tasks were high in cognitive demand at the selection and planning phases. These findings contrast those of the 1996 TIMSS study and provide some evidence that the incentives of NCTM and others are taking hold in specific contexts. Specifically, in situations where high school mathematics teachers are involved in task-focused professional development and provided with high-quality curricular materials, there is strong potential for them to successfully implement tasks at high levels. The findings of the present study do not necessarily generalize across similar contexts. However, it appears that the MMR and ATC PD programs fostered environments where high-level task implementation is possible. Similar PD programs may also provide such promise, though additional research is necessary to explore such programs.

Focusing on Active Student Engagement

A consistent finding among the themes for teachers' task selection, planning, setup, and implementation was a focus on student engagement with a task. In the present study, the participating teachers identified active student engagement as a motivating factor for selecting their tasks. When planning, the teachers established learning goals for their students and anticipated how their students might interact with the mathematics and with each other. During task setup, the teachers engaged their students in relevant preliminary discussions and set expectations for students' subsequent work on a task. Finally, during implementation, the teachers took action to monitor and facilitate students' engagement as it occurred. The focus on student engagement at each of the four task phases suggests that teachers were concerned with student engagement throughout their use of tasks. Before, during, and even after their instruction, the teachers attended to how their students' potential and actual engagement and how such engagement contributed to their students' mathematical learning.

Teachers' focus on active student engagement is critical because student engagement influences the cognitive demand of a task during task setup and task implementation. According to Henningsen and Stein (1997), students learn to "do mathematics" when they "engage actively in rich, worthwhile mathematical activity" (p. 524). Moreover, NCTM (2014) advocates for active, student-centered instructional practices because these approaches can enhance students' ability to solve problems and to reason mathematically. Active student engagement alone is not enough to maintain the high cognitive demand of a task (Stein et al., 2009); however, it is one of the essential ingredients that can promote high-level student thinking and reasoning.

A key feature of the MMR course is its student-centered nature. By design, rather than lecturing, MMR teachers are expected to pose rich modeling problems and support students' engagement by encouraging productive struggle and challenging students to identify their own solution pathways. In keeping with this design, the present study found, for the most part, that the MMR teachers were effective at implementing the course, and the PD was effective at preparing teachers to do so. Though the participating teachers' understanding of cognitive demand was still developing in some ways, they consistently planned for and taught engaging lessons, during which students explored real-world situations and developed their own problem-solving strategies. Even in cases where the cognitive demand of a task declined during implementation, the students were still engaged in some form of work and thinking (e.g., collecting data and generating mathematical models using technology). In other cases, student engagement contributed to the high cognitive demand of tasks that were set up and implemented. In general, the ways in which teachers engaged their students aligned with the goals of the MMR course.

Research Using the Instructional Quality Assessment and the Task Analysis Guide

The present study incorporated the IQA (Boston, 2012) and the TAG (Stein et al., 2009), two rubrics for analyzing the cognitive demand of a task. These are not the only two tools that can be used to determine the rigor of mathematics instruction and the potential for student learning, but they were chosen because of their clear focus on instructional tasks and cognitive demand. Moreover, the IQA rubrics were designed to closely align with the TAG and measure similar attributes of a task. As the researcher, I worked with another trained observer to use the IQA rubrics because they were designed for use by two observing researchers who would then reach a consensus on the rating of the cognitive demand of a task. I had research participants use the TAG because they were trained to use to use it during PD but were not trained to use the IQA.

As depicted in Figure 3 on page 104, the TAG and the IQA are comparable at their three lowest levels: *nonmathematical activity* according to the TAG warrants an IQA score of 0, *memorization* tasks in the TAG score a 1 on the IQA, and *procedures without connections* tasks in the TAG score a 2 on the IQA. That is, low cognitive demand tasks are analyzed in essentially the same way by either the TAG or the IQA. However, the categories for high cognitive demand tasks do not translate as directly between the two rubrics: both *procedures with connections* and *doing mathematics* tasks have the potential to score either 3 or 4 on the IQA. The difference between IQA scores of 3 and 4 is explicit evidence of students' mathematical thinking and reasoning, whereas the difference between *procedures with connections* and *doing mathematics* tasks is altogether different in the TAG. The IQA also focuses on the use of accountable talk moves, especially in the Expanded IQA Task Setup Rubrics, whereas the TAG does not.

By comparing teachers' analysis of tasks using the TAG with the researcher's using the IQA, the present study has highlighted the nuances between the TAG and the IQA regarding high cognitive demand tasks. In some instances, it might be more logical for the researcher to use the TAG alongside teachers, and in other cases, both parties might be trained to use the IQA rubrics. Candela, for example, has conducted PD studies in which teachers have learned about the IQA and have used it to analyze and reflect on their own instructional practices (e.g., Candela, 2016; Candela & Boston, 2019). Such studies have explored the use of the IQA as a "professional learning tool" (Candela & Boston, 2019, p. 530) to support mathematics teachers' growth in selecting and implementing high level tasks. Though it may still be worthwhile to use and compare both the TAG and the IQA in the same study, researchers must be intentional when determining which to use and when to use it.

Recommendations for Future Practice

Based on the findings of present study, I have identified three recommendations for future practice. The first is that the use of high-quality curricular materials contributed to teachers' ability to set up and implement tasks maintaining high cognitive demand; therefore, the present study highlights the importance of providing teachers with such materials for other high school mathematics courses in addition to MMR. The second recommendation suggests ways in which teachers can maintain the cognitive demand of tasks based on the findings of the present study. The third recommendation focuses on teacher professional development, particularly PD incorporating the Stein et al. (2009) MTF and TAG.

Curricular Materials

Trends in teachers' task use and in IQA scores highlight the importance of rigorous curricular materials in school mathematics. One of the trends that was identified using the IQA was that teachers' tasks consistently scored 3 and 4 for Phases 1 and 2. This finding suggests that the tasks were all high in cognitive demand for both phases. Though additional research is needed to determine whether all the MMR tasks are high in cognitive demand, the sample of tasks in the present study suggest that many others might be. Stein and colleagues (2009) conjectured that selecting cognitively demanding tasks "appears to be a necessary condition" (p. 5) for engaging students in high-level thinking and reasoning; evidence from the present study shows that, for teachers involved in PD focusing on the cognitive demand of mathematical tasks, those who select high cognitive demand tasks are likely to implement the tasks in such ways, though 3 of the 8 MMR teachers failed to do so for various reasons. However, the use of rigorous curricular materials surely influenced the success of the remaining 5 teachers.

Another trend was that most of the teachers made few significant changes to their tasks when planning, aside from those that might make them more suitable for remote learning and safer for face-to-face instruction during the COVID-19 pandemic. There is not enough evidence in the interview data to support the following claim, but I suspect that the participating MMR teachers chose not to make substantial changes to other aspects of their tasks because they were strong tasks to begin with. Some of the teachers indicated their stance that the MMR materials were "effective" and that they would use them to teach similar content in other courses. Teachers who are provided with such materials can spend more of their planning time focused on pedagogy and anticipating how they might support and enhance student learning, rather than spending time searching for and adapting materials to fit their needs.

The inclusion of lesson plans with the MMR tasks may also have supported teachers' task setup and implementation in the present study. The MMR teachers attended to the lesson plans for their tasks and included the state-recommended instructional protocols when they taught, including the use of videos and related whole-class discussions. Teachers set up the tasks in ways that were suggested by the MMR lesson plans and practiced instructional strategies that were recommended during task implementation. The provided lesson plans may have supported high-level task setup and implementation and benefited both teachers and their students. Therefore, the availability of rigorous instructional materials cannot be overstated; high school mathematics in the United States would significantly benefit if such materials were made widely available for not only MMR but more standard courses such as Algebra, Geometry, and Precalculus.

Maintaining Cognitive Demand in Practice

The present study resonates with previous research, providing further evidence that selecting a high cognitive demand task does not necessarily guarantee high-level implementation. The present study also highlights the significance of the implementation phase, the task phase where decline in cognitive demand is most common. To support students' conceptual understanding and mathematical sensemaking, more teachers should be provided with opportunities to discover and explore (a) the Stein et al. (2009) MTF, (b) the TAG, (c) and the four-phase framework that I presented in Figure 1 on page 50. These tools are essential if teachers are to understand the complexity of mathematical tasks and their influence on student learning. Teachers can learn about these ideas in PD, at mathematics education conferences such as those sponsored by the NCTM, its affiliates, and other similar organizations in the United States and abroad, and through teacher education programs at colleges and universities. By doing so, mathematics teachers can begin to develop mindsets focused on cognitive demand that enhance their use of tasks across the four phases.

The setup and implementation phases are less predictable because of potential student interaction and require attention so that the cognitive demand of a task does not decline. Teachers who support their students in preliminary discussions of the real-world features and mathematical relationships necessary to engage with a task can enhance student learning and productivity (Jackson et al., 2012, 2013); however, teachers must also be conscientious to avoid doing the work of the task *for* their students and thus reducing the cognitive demand. Evidence from the present study suggests that mathematics teachers can help their students visualize aspects of a task in other ways by asking them to think of alternative ways to represent the situation being investigated. Therefore, without lowering the cognitive demand, teachers can suggest alternative representations for students and still let the students do the work on their own.

Alternatively, teachers may maintain and even enhance the cognitive demand of a task by pressing students to explain their thinking and reasoning. Tasks that require such evidence of students' understanding are crucial, but the teacher can further promote this by establishing classroom environments where students are required to provide explanations regularly. The cases of Gwen and Henry provide evidence that it is possible to promote such an environment despite the use of tasks that do not require explicit evidence of students' thinking and reasoning. Enhancing the rigor of a task is a skill that may not often be addressed in teacher education programs and teacher PD; therefore, the present study highlights the need to support mathematics teachers in this area. Teachers who are not provided with high-quality instructional materials may find it necessary to modify tasks so that they can engage their students in high-level thinking and reasoning frequently and consistently. As in the case of Gwen, however, setting the expectation for high-level thinking and reasoning is not enough. Enhancing the quality of a task is a beneficial skill for teachers to learn, but they must also learn to hold students accountable for providing mathematical evidence and justification with their responses.

Though some of the teachers in the present study communicated their frustration with remote teaching and learning, high cognitive demand tasks may also thrive in such spaces. Teachers who maintained high cognitive demand with their tasks used technology appropriately to engage their students, to monitor and facilitate their work, and to promote discussions between student groups and among the whole class. The use of technology such as Pear Deck and Nearpod, a similar alternative, allows teachers to observe students' work in real-time. This enables teachers to provide nearly instantaneous feedback and formative assessment aimed at enhancing student learning. A crucial element of remote instruction that is conducive to maintaining high cognitive demand is that students frequently interact with each other and with the mathematics; though it can be challenging to get students to talk in online environments for various reasons, teachers must provide opportunities for them to do so because it allows them to share and build on each other's insights.

Professional Development

The MMR PD provided teachers with an enormous amount of information: The teachers learned about mathematical tasks, cognitive demand, and how to analyze both the potential (Phase 1: Selection) and the implementation (Phase 4: Implementation) of a task using the TAG. In addition, they were introduced to new curricular content and instructional materials as well as innovative pedagogical strategies. And in the case of the 2020–2021 school year, all of this was done amidst the unprecedented challenges presented by the COVID-19 pandemic. Therefore, it is understandable that that the TAG was not the teachers' first and foremost priority, nor might it have been for the MMR PD leaders. Fortunately, the MMR PD not only included 3 days of summer workshops but also regular follow-up meetings throughout the ensuing school year. However, in non-COVID years, the summer PD has been 4 days in length and has been face to face.

The MMR PD and similar PD programs should include the MTF, the TAG, the IQA, and a study of cognitive demand and the PD leaders should be sure to allocate sufficient time for teachers to engage deeply and meaningfully in (a) working through mathematical tasks; (b) studying the MTF, the TAG, the IQA, and the associated levels of cognitive demand; and (c) analyzing instructional tasks using these tools. Moreover, PD coupled with instructional materials containing high cognitive demand tasks enhance teachers' ability to integrate such tasks into their regular teaching practices, and the PD support helps teachers to maintain high cognitive demand throughout the task's lifespan.

The MMR PD provided teachers with high-quality curricular materials in addition to what they learned about mathematical tasks. These materials surely influenced the number of tasks that scored 3 and 4 on the IQA across the four task phases. Though the teachers involved in the present study provided generally strong analyses of tasks in Phases 1–2, assessing the cognitive demands of tasks in Phases 3–4 was more challenging, especially for first-year MMR teachers. Moreover, the teachers were not accustomed to analyzing the cognitive demands of tasks using four separate task phases; therefore, future task-focused PD should provide specific attention to (a) the cognitive demand of a task at each of the four task phases and (b) potential changes to a task between each phase and how such changes influence the cognitive demand of the task. Teachers need opportunities to analyze the cognitive demand of tasks at each phase but should also have a chance to select, plan, set up, and implement their own tasks, with opportunities to reflect on and discuss the changing cognitive demand with other teachers. Such PD will not be successful in a single setting over the course of a day or a weekend; for teachers to be able to thoroughly reflect on their practice with each other, a PD program should include several follow-up sessions throughout the course of a semester or year.

Desimone (2002, 2011) argued that five features influence the effectiveness of PD and inspire changes in teachers' knowledge and practice. The first feature, emphasis on content and how students learn, is one of the features of the MMR PD. The MMR PD focuses primarily on the content of the MMR course and the use of student-centered pedagogical approaches. The second of Desimone's features of effective PD is opportunities for teachers to participate and learn actively; in both the 2019 and 2020 summer MMR workshops, PD leaders engaged teachers in working through some of the MMR tasks so that they could learn by *doing* as their students should. The third feature, opportunities for teachers to act and serve as leaders, occurred as MMR teachers provided feedback on the course and served on curriculum development and PD committees.

The MMR PD was ongoing, providing teachers with continuous learning opportunities through regular follow-up meetings during each school year; Desimone's fourth feature is the length of PD programs, asserting that one-day and weekend-long programs are less effective at promoting long-term teacher growth than programs that span a semester or year. Finally, the fifth feature is collective participation from teachers who share common characteristics. This was achieved by the MMR PD through the recruitment of high school mathematics teachers teaching the same course, MMR, and the development of a statewide learning community that met regularly to discuss lessons and share resources with each other. In summary, the MMR PD possessed many high-quality features. The findings of the present study further suggest that the MMR PD has positively influenced teachers' practice through their use of instructional tasks. Future PD programs should strive to adopt these evidence-based approaches.

Recommendations for Future Research

The following recommendations for future research serve to extend the present study and address its limitations. As the researcher, I have reflected on my methods and can provide recommendations that may yield more rich data and address the limitations of the study. Because this study was only a short-term glimpse into teachers' perspectives, longitudinal studies are necessary to explore changes in teachers' perspectives over time. In addition, this study was delimited to the context of two task-focused PD programs in Ohio; research exploring additional PD programs and teachers' typical practice (outside the context of PD) is necessary to determine how the findings differ in different contexts. Finally, more research should explore teachers' understanding of the Stein et al. (2009) levels of cognitive demand and the four task phases to enhance teacher preparation and training programs.

Enhancing Similar, Future Studies

As I reflect on the present study, I realize that the research design and methods were limited to my capability as a researcher (Glesne, 2016; Patton, 2015). As my interviews with teachers progressed from task selection to planning, setup, and implementation, the dialogue tended to focus more on the *actions* of teachers than on the *reasons* for such actions. To enhance the potential for gathering rich data, researchers conducting similar studies might enhance their methods in two ways: first, time constraints might have restricted my ability to delve deeper into teachers' rationales than I had hoped. Discussing, in depth, two phases of teachers' task use and the cognitive demand of a task at both phases might be more than can be accomplished within a single hour. Second, my inexperience as a researcher might have limited my ability to probe and converse with teachers during our interviews. Therefore, in similar, future studies, I recommend allocating more time for each of the four task phases and developing deeper probing questions so that teachers might be pressed to explain their reasoning in greater detail. These two suggestions can enhance the richness of the data obtained in future studies and lead to additional research findings.

Longitudinal Studies

One of the limitations of the present study is the timeframe for which it was set. Two observations of a teacher's practice, with one interview before and one afterward, provides only a snapshot of each participant's task use. Therefore, future research of a similar nature should explore teachers' perspectives and use of mathematical tasks prior to, during, and following the conclusion of a PD program. Longitudinal studies would extend beyond a single snapshot and provide insights regarding the way that PD may influence teachers' perspectives and, hopefully, enhance the ways that they use instructional tasks in their teaching. Previous research has focused on the latter but has not included the former, nor has it addressed the way in which teachers' changing perspectives influence their changing task use. The Ohio MMR pilot course serves as a facilitator through which long-term studies may continue and extend the work of this dissertation.

Differing Contexts

The present study is delimited to the context of two PD programs facilitated within the state of Ohio. Moreover, the qualitative nature of the research does not lend itself to be generalizable to teachers in other contexts. Additional research is necessary to explore mathematics teachers' perspectives and their use of tasks in various contexts, including (a) teachers in similar PD programs in different parts of the world, (b) teachers in Ohio involved in similar PD programs, and (c) teachers in Grades PreK–8 and teachers of courses other than MMR. Conducting similar studies in these differing contexts may provide insights that were not obtained through the present study. In addition, such studies might support the findings of the present study and further contribute to the body of literature on mathematical tasks. Because mathematics teaching is cultural (Stigler & Hiebert, 1999), conducting research similar to the present study in another country (e.g., Germany or Japan) would likely introduce new variables based on the interactions between teachers and students. Teachers who collaborate frequently and purposefully through lesson studies, such as those in countries like Japan, might reflect on their instruction in ways that teachers in the United States did not in the present study.

Additional research is also needed with teachers who are not currently involved in PD and those who are not teaching pilot courses. One of the potential reasons that the MMR teachers made few substantial changes to their tasks when planning is because the materials were provided as a part of the PD for the pilot course. Some teachers acknowledged the quality of the course materials whereas others acknowledged ODE's desire for "consistency" when implementing the lessons because of their research. I suspect that teachers who are neither involved with PD nor teaching a pilot course may select tasks from a variety of sources (though textbooks may be a predominant resource) and may adapt them in more substantial ways than those participating in the present study. The single non-MMR teacher, Isabel, provides a single instance of a teacher who found and adapted a task from the internet.

One of the delimitations of the present study was that I required teachers to use what they considered to be a high cognitive demand task. What might have resulted if the same teachers were not given such a requirement? Some of the teachers might have selected a low cognitive demand task, which would certainly have led to other findings and conclusions. What if the teachers had been required to select a low cognitive demand task and then modify it to be a high-level task? These types of open questions require additional research to explore.

The present study was also conducted amidst the COVID-19 pandemic, requiring numerous adaptations that influenced data collection. All the data for the present study were collected online, whereas research in a typical setting would have been done face-to-face; future research in a more conventional setting may yield additional findings. Conducting the present study in a typical face-to-face setting might have led to different findings because teachers' use of tasks and interactions with students might have been different.

Teachers' Understanding of Cognitive Demand

The present study found evidence indicating that high school mathematics teachers perceive cognitive demand in various ways and at differing levels of understanding, despite their participation in the same PD program. Some of the participating teachers in this study had engaged in the MMR PD, ATC PD, and possibly others involving the TAG numerous times; such teachers are expected to have developed a deeper knowledge of how to analyze tasks based on their cognitive demands. However, those with similar levels of experience with the TAG conceptualized cognitive demand in different ways. Therefore, further research is needed to explore the ways in which teachers develop understanding of cognitive demand as a theoretical construct and how mathematical tasks are analyzed using the TAG. Researchers such as Arbaugh and Brown (2005) and Boston and Smith (2009) have investigated the influence of PD in improving high school mathematics teachers' understanding of tasks and the TAG but have not done so according to the four task phases of selecting, planning, setting up, and implementing. Additional research is necessary to address teachers' growth as they learn to analyze tasks at all four phases. Future studies of this nature will enhance the effectiveness of future PD programs and may enhance preservice teacher education programs as well.

Future research could also explore teachers' analysis of tasks with similar features but different levels of cognitive demand. One study of interest might require high school mathematics teachers to compare two tasks with similar superficial features (e.g., including a real-world context or involving graphing), one with high cognitive demand and one with low cognitive demand. Teachers' analysis of the tasks and their perspectives of how the tasks compare might yield further insights related to those of the present study. Similarly, researchers might also present teachers with low cognitive demand tasks and prompt them to modify the tasks in ways that enhance their cognitive demand. Such a study could explore the reasons why teachers make the modifications that they do and how the teachers think that the changes will enhance the cognitive demand of the task.

Summary and Conclusion

The purpose of this study was to explore high school mathematics teachers' perspectives as they progressed through the phases of selecting, planning, setting up, and implementing high cognitive demand tasks. Past research has explored teachers' task use using quantitative approaches, indicating that the cognitive demand of a task tends to decline from the first phase to the last. However, few studies had focused on teachers' perspectives, and none had incorporated a four-phase model for the progression of a task. Therefore, research investigating teachers' perspectives regarding mathematical tasks and cognitive demand is crucial to advance the field of mathematics teacher education and to support high school mathematics teachers as they endeavor in this challenging work.

By conducting interviews, observing classroom instruction, and analyzing instructional tasks and student work, I have gathered evidence to describe 9 high school mathematics teachers' use of instructional tasks. In addition, I have discovered their perspectives concerning how they select, plan, set up, and implement such tasks in their teaching and how they conceptualize the cognitive demand of a task from one phase to the next. The cross-case thematic analysis of teachers' interviews led to the establishment of 18 themes split among the four task phases: generally, the teachers selected tasks that were suitable for their respective learning environments (e.g., face-to-face, remote, or hybrid), encouraged students to engage actively, addressed important content and skills, and included genuine real-world contexts. Participating teachers were flexible with their planning but attended to goals and objectives specified by ODE in the MMR lesson plans and their own goals for student learning. They also anticipated ways in which their students might engage and struggle with their tasks and adjusted tasks according to their teaching environments.

Most of the participating teachers included some form of preliminary discussion prior to students' engagement with the task, though some had their students begin working on their tasks right away. The teachers typically set expectations for students' engagement and, when relevant, had their students brainstorm by asking questions such as "What do you notice?" and "What do you wonder?" When implementing the tasks, teachers encouraged productive struggle as they facilitated and monitored their progress. However, they intervened when appropriate by asking questions, providing support, and eliciting evidence of students' thinking and reasoning throughout the process.

The analysis of teachers' tasks using the IQA showed that some teachers set up and implemented their tasks in ways that maintained high-level cognitive demands. However, other tasks declined in cognitive demand either because (a) students did not explain their mathematical thinking and reasoning or (b) the emphasis of the task shifted toward routinized procedures and finding answers. Despite previous research findings suggesting that increases in cognitive demand are rare, teachers in the present study increased the cognitive demand of tasks by setting the expectation that students should explain their thinking and by pressing for explanations. In both instances, student explanations were not called for by the original task. Mismatches between researchers' and teachers' analysis of task cognitive demands occurred only in task implementation; such instances occurred when (a) teachers acknowledged some decline in cognitive demand but did not adjust their TAG classifications and when (b) a teacher communicated a lack of understanding of the levels of cognitive demand and focused on nonmathematical evidence of students' thinking and reasoning.

The findings of this study suggest that the PD and associated resources provided to the MMR teachers substantially influenced their use of mathematical tasks. This is not only evident in the tasks they selected, but also in the teachers' attention to the goals and lesson plans developed by ODE. Moreover, the participating MMR teachers set up and implemented tasks as suggested in the ODE lesson plans, including instructional segments (e.g., brainstorming and data-gathering segments) and recommended pedagogical strategies (e.g., asking questions and eliciting evidence of students' thinking). The effectiveness of the MMR PD seems to have influenced the high-level task potential that was evident across the teachers participating in the present study; that is, the PD and resources provided by ODE enabled the teachers to select tasks with high cognitive demand and maintain the cognitive demand into the planning phase.

Task setup across the 8 MMR teachers was generally consistent with the recommendations provided in the ODE lesson plans. All the MMR tasks maintained their potential to engage students in cognitively demanding work and thinking. Though 5 of the 8 MMR tasks were implemented at high cognitive levels, 3 declined into the low cognitive domain of *procedures without connections*. Moreover, the teachers who implemented these three tasks failed to notice the decline into lower cognitive activity. This suggests that the MMR PD and similar PD programs should emphasize task implementation and teacher strategies to maintain high-level cognitive demand. In addition, PD programs should educate participants about the four phases of selecting, planning, setting up, and implementing tasks and the various ways that cognitive demand may change from one phase to the next. Enhancing teachers' understanding of cognitive demand and the complex nature of mathematical tasks will result in an increase in the rigor of tasks that their students engage with; therefore, doing so will enhance student learning and understanding in school mathematics.

References

ACT. (2013). *ACT 2013 profile report: Graduating class of 2013—National.* Author.

Ambrose, R. (2004). Initiating change in prospective elementary school teachers' orientations to mathematics teaching by building on beliefs. *Journal of Mathematics Teacher Education, 7*, 91–119. http://dx.doi.org/10.1023/B:JMTE.0000021879.74957.63

Arbaugh, F., & Brown, C. A. (2005). Analyzing mathematical tasks: A catalyst for change? *Journal of Mathematics Teacher Education*, *8*(6), 499–536. http://dx.doi.org/10.1007/s10857-006-6585-3

Ball, D. L, & Forzani, F. M. (2010). Teaching skillful teaching. *Educational Leadership, 68*(4), 40–45.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59*, 389–407. http://dx.doi.org/10.1177/0022487108324554

Beswick, K. (2011). Teachers' beliefs about school mathematics and mathematicians' mathematics and their relationship to practice. *Educational Studies in Mathematics, 79*, 127–147. http://dx.doi.org/10.1007/s10649-011-9333-2

Boaler, J., & Staples, M. (2008). Creating mathematical futures through an equitable teaching approach: The case of Railside School. *Teachers College Record, 110,* 608–645.

Borg, W. R. (1980). Time and school learning. In C. Denham & A., Lieberman (Eds.), *Time to learn*. National Institute of Education.

Boston, M. D. (2006). *Developing secondary mathematics teachers' knowledge of and capacity to implement instructional tasks with high level cognitive demands* [Doctoral dissertation, University of Pittsburgh].

Boston, M. D. (2012). Assessing the quality of mathematics instruction. *Elementary School Journal, 113,* 76–104. http://dx.doi.org/10.1086/666387

Boston, M., Bostic, J., Lesseig, K., & Sherman, M. (2015). A comparison of mathematics classroom observation protocols. *Mathematics Teacher Educator, 3,* 154–175. http://dx.doi.org/10.1007/s10857-019-09445-0

Boston, M. D., & Candela, A. G. (2018). The instructional quality assessment as a tool for reflecting on instructional practice. *ZDM Mathematics Education, 50,* 427–444. http://dx.doi.org/10.1007/s11858-018-0916-6

Boston, M. D., & Smith, M. S. (2009). Transforming secondary mathematics teaching : Increasing the cognitive demands of instructional tasks used in teachers' classrooms. *Journal for Research in Mathematics Education*, *40*(2), 119–156.

Boston, M. D., & Smith, M. S. (2011). A 'task-centric' approach to professional development: Enhancing and sustaining mathematics teachers' ability to implement cognitively challenging mathematical tasks. *ZDM Mathematics Education, 47*(6–7), 965–977. http://dx.doi.org/10.1007/s11858-011-0353-2

Boston, M. D., & Wilhelm, A. G. (2017). Middle school mathematics instruction in instructionally focused urban districts. *Urban Education, 52*(7), 829–861. http://dx.doi.org/10.1177/0042085915574528

Boston, M., & Wolf, M. K. (2006). Assessing academic rigor in mathematics instruction: The development of the instructional quality assessment toolkit. National Center for Research on Evaluation, Standards, and Student Testing (CRESST) Report No. 672. http://dx.doi.org/10.1037/e644922011-001

Brinkmann, S. & Kvale, S. (2015). *InterViews: Learning the craft of qualitative research interviewing* (3rd ed.). Sage.

Brophy, J. E. (1979). Teacher behavior and its effects. *Journal of Educational Psychology, 71*(6), 733–750. http://dx.doi.org/10.1037/0022-0663.71.6.733

Candela, A. G. (2016, November 3–6). *Mathematics teachers' perspectives on factors affecting the implementation of high cognitive demand tasks* [Paper presentation]. North American Chapter of the International Group for the Psychology of Mathematics Education 38th Annual Meeting, Tuscon, AZ, United States.

Candela, A. G., & Boston, M. (2019, November 14–17). *Mathematics teachers' perceptions of the Instructional Quality Assessment as a reflective tool* [Paper presentation]. North American Chapter of the International Group for the Psychology of Mathematics Education 41st Annual Meeting, St Louis, MO, United States.

Chapman, O. (2013). Mathematical-task knowledge for teaching. *Journal of Mathematics Teacher Education, 16*(1), 1–6. http://dx.doi.org/10.1007/s10857-013-9234-7

Charalambous, C. Y. (2010). Mathematical knowledge for teaching and task unfolding: An exploratory study. *Elementary School Journal, 110*(3), 247–278. http://dx.doi.org/10.1086/648978

Clare, L., & Aschbacher, P. (2001). Exploring the technical quality of using assignments and student work as indicators of classroom practice. *Educational Assessment, 7*(1), 39–59. http://dx.doi.org/10.1207/S15326977EA0701_5

Clement, L. L. (1999). The constitution of teachers' orientations toward teaching mathematics. (Doctoral dissertation, University of California at San Diego and San Diego State University, 1999). *Dissertation Abstracts International, 60*(06), 1949A.

Cobb, P., Jackson, K., Henrick, E., & Smith, T. M. (2018). Investigating and supporting instructional improvement. In P. Cobb, K. Jackson, E. Henrick, T. M. Smith, & the MIST team (Eds.), *Systems for instructional improvement: Creating coherence from the classroom to the district office* (pp. 1–14). Harvard Education Press.

Cobb, P., Wood, T., & Yackel, E. (1990). Chapter 9: Classrooms as learning environments for teachers and researchers. *Journal for Research in Mathematics Education. Monograph, 4,* 125–210. http://dx.doi.org/10.2307/749917

College Board. (2013). *2013 SAT Report on College and Career Readiness.* College Board. http://dx.doi.org/10.1119/1.4830084

Collopy, R. (2003). Curriculum materials as a professional development tool: How a mathematics textbook affected two teachers' learning. *Elementary School Journal, 103*, 287–311. http://dx.doi.org/10.1086/499727

Cooney, T. J. (1999). Conceptualizing teachers' ways of knowing. *Educational Studies in Mathematics, 38*, 163–187. http://dx.doi.org/10.1007/978-94-017-1584-3_8

Davis, E. A. (2006). Preservice elementary teachers' critique of instructional materials for science. *Science Education, 90*(2), 348–375. http://dx.doi.org/10.1002/sce.20110

Davis, R. B., & McKnight, C. (1976). Conceptual, heuristic, and S-algorithmic approaches in mathematics teaching. *Journal of Children's Mathematical Behavior, 1*(Supplement 1), 271–286.

Desimone, L. M., Porter, A. C., Garet, M. S., Yoon, K. S., & Birman, B. F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Educational Evaluation and Policy Analysis, 24*(2), 81–112. http://dx.doi.org/10.3102/01623737024002081

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher, 38*(3), 181–199. http://dx.doi.org/10.3102/0013189X08331140

Desimone, L. M. (2011). A primer on effective professional development. *Kappan, 92*(6), 68–71. http://dx.doi.org/10.1177/003172171109200616

Desmos. (2021). *Point collector: Lines.* https://teacher.desmos.com/activitybuilder/custom/57e55435920c2cd506168126

Doyle, W. (1977). Paradigms for research on teacher effectiveness. *Review of Research in Education, 5*(1), 163–198. http://dx.doi.org/10.2307/1167174

Doyle, W. (1983). Academic work. *Review of Educational Research, 53*(2), 159–199. http://dx.doi.org/10.3102/00346543053002159

Doyle, W. (1988). Work in mathematics classes: The context of students' thinking during instruction. *Educational Psychologist, 23*(2), 167–180. http://dx.doi.org/10.1207/s15326985ep2302_6

Earnest, D., & Amador, J. M. (2019). Lesson planimation: Prospective elementary teachers' interactions with mathematics curricula. *Journal of Mathematics Teacher Education, 22*, 37–68. http://dx.doi.org/10.1007/s10857-017-9374-2

Fieman-Nemser, S., McDiarmid, G.W., Melnick, S., & Parker, M. (1988). *Changing beginning teachers' conceptions: A study of an introductory teacher education course.* National Center for Research on Teacher Education and Department of Teacher Education, Michigan State University.

Foley, G. D., Strayer, J. F., & Regan, B. (2010). High school mathematics teacher professional development in data analysis, probability, and statistics. *Journal of the Research Center for Educational Technology, 6*(2), 4–17 and three appendices. http://rcetj.org/index.php/rcetj/article/viewArticle/132

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal, 38*(4), 915–945. http://dx.doi.org/10.3102/00028312038004915

Glesne, C. (2016). *Becoming qualitative researchers: An introduction* (5th ed.). Longman.

Gonzales, P., Williams, T., Jocelyn, L., Roey, S., Kastberg, D., & Brenwald, S. (2008). *Highlights from TIMSS 2007: Mathematics and science achievement of U.S.*

fourth- and eighth-grade students in an international context. National Center for Education Statistics. https://files.eric.ed.gov/fulltext/ED503625.pdf

Good, T. (1979). Teacher effectiveness in the elementary school. *Journal of Teacher Education, 30*(2), 52–64. http://dx.doi.org/10.1177/002248717903000220

Graves, D. H. (1975). Research doesn't have to be boring. *Language Arts, 56*, 76–80.

Green, T. F. (1971). *The activities of teaching.* McGraw-Hill.

Grouws, D. A., Tarr, J. E., Chávez, Ó., Sears, R., Soria, V. M., & Taylan, R. D. (2013). Curriculum and implementation effects on high school students' mathematics learning from curricula representing subject-specific and integrated content organizations. *Journal for Research in Mathematics Education*, *44*(2), 416–463. http://dx.doi.org/10.5951/jresematheduc.44.2.0416

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods, 18*(a), 59–82. http://dx.doi.org/10.1177/1525822X05279903

Guskey, T. R. (2002). Professional development and teacher change. *Teachers and Teaching: Theory and Practice, 8*(3), 381–391. http://dx.doi.org/10.1080/135406002100000512

Henningsen, M., & Stein, M. K. (1997). Mathematical tasks and student cognition: Classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education*, *28*(5), 524–549. http://dx.doi.org/10.2307/749690

Heyd-Metzuyanim, E., Smith, M., Bill, V., & Resnick, L. B. (2019). From ritual to explorative participation in discourse-rich instructional practices: A case study of teacher learning through professional development. *Educational Studies in Mathematics, 101*, 273–289. http://dx.doi.org/10.1007/s10649-018-9849-9

Hiebert, J., & Wearne, D. (1993). Instructional tasks, classroom discourse, and students' learning in second-grade arithmetic. *American Educational Research Journal, 30*(2), 393–425. http://dx.doi.org/10.3102/00028312030002393

Hill, H. C., & Ball, D. L. (2004). Learning mathematics for teaching: Results from California's mathematics professional development institutes. *Journal of Research in Mathematics Education, 35*(5), 330–351. http://dx.doi.org/10.2307/30034819

Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education, 39*(4), 372–400.

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

Hill, H., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42*(2), 371–406. http://dx.doi.org/10.3102/00028312042002371

Jackson, K., Garrison, A., Wilson, J., Gibbons, L., & Shahan, E. (2013). Exploring relationships between setting up complex tasks and opportunities to learn in concluding whole-class discussions in middle-grades mathematics instruction. *Journal for Research in Mathematics Education, 44*(4), 646–682. http://dx.doi.org/10.5951/jresematheduc.44.4.0646

Jackson, K. J., Shahan, E. C., Gibbons, L. K., & Cobb, P. A. (2012). Launching complex tasks: Consider four important elements of setting up challenging mathematics problems to support all students' learning. *Mathematics Teaching in the Middle School, 18*(1), 24–29. https://doi.org/10.5951/mathteacmiddscho.18.1.fm

James, C., Philiben, L., & Knievel, M. (2016). Doing the math: Supporting student justifications. *Mathematics Teaching in the Middle School, 21*(7), 416–423. https://doi.org/10.5951/mathteacmiddscho.21.7.0416

Johnson, M. (Ed.). (1980). *Toward adolescence: The middle school years.* 79th Yearbook of the National Society for the Study of Education (Part 1). University of Chicago Press.

Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics.* National Academy Press. https://doi.org/10.17226/9822

King, A. (2015). Ramp it up. *Mathematics Teaching in the Middle School, 21*(5), 320. https://doi.org/10.5951/MATHTEACMIDDSCHO.21.5.0320

Knoll, E., Ernest, P., & Morgan, S. (2004). Experiencing research practice in pure mathematics in a teacher training context. In M. J. Hoines & A. B. Fuglestad (Eds.), *Proceedings of the 28th annual conference of the International Group for* *the Psychology of Mathematics Education* (Vol. 3, pp. 161–168). Bergen University College.

Lee, H. S., Coomes, J., & Yim, J. (2019). Teachers' conceptions of prior knowledge and the potential of a task in teaching practice. *Journal of Mathematics Teacher Education, 22*(2), 129–151. http://dx.doi.org/10.1007/s10857-017-9378-y

Lincoln, Y. S., & Guba, E. G. (1985). Paradigmatic controversies, contradictions, and emerging confluences. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp.163–188). Sage.

MacKay, R. (1978). How teachers know: A case of epistemological conflict. *Sociology of Education, 51*, 177–187. http://dx.doi.org/10.2307/2112663

Matsumura, L. C., Garnier, H., E., Pascal, J., & Valdes, R. (2002). *Measuring instructional quality in accountability systems: Classroom assignments and student achievement.* National Center for Research on Evaluation, Standards, and Student Testing (CRESST) Report No. 582. http://dx.doi.org/10.1207/S15326977EA0803_01

Matsumura, L. C., Garnier, H., Slater, S. C., & Boston, M. (2008). Toward measuring instructional interactions 'at-scale.'. *Educational Assessment, 13,* 267–300. http://dx.doi.org/10.1080/10627190802602541

Matsumura, L. C., Slater, S. C., Junker, B., Peterson, M., Boston, M., Steele, M., & Resnick, L. (2006). Measuring reading comprehension and mathematics instruction in urban middle schools: a pilot study of the instructional quality assessment. National Center for Research on Evaluation, Standards, and Student Testing (CRESST) Report #681.

Mehan, H. (1974). Accomplishing classroom lessons. In A. V. Cicourel, K. H. Jennings, S. H. M. Jennings, K. C. W. Leiter, R. MacKay, J. Mehan, & D. Roth (Eds.), *Language use and school performance.* Academic Press. http://dx.doi.org/10.1016/B978-0-12-174950-7.50007-4

Mehan, H. (1979). *Learning Lessons: Social Organization in the Classroom.* Harvard University Press. http://dx.doi.org/10.4159/harvard.9780674420106

Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation.* Wiley.

Monk, D. H. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of Education Review, 13*(2), 125–145. http://dx.doi.org/10.1016/0272-7757(94)90003-5

Mullis, I. V. S., Martin, M. O., Foy, P., Olson, J. F., Preuschoff, C., Erberber, E., Arora, A., & Galia, J. (2008). *TIMSS 2007 international mathematics report: Findings from IEA's trends in international mathematics and science study at the fourth and eighth grades.* TIMSS & PIRLS International Study Center, Boston College.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Author.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Author.

National Council of Teachers of Mathematics. (2014). *Principles to actions: Ensuring mathematical success for all*. Author.

National Council of Teachers of Mathematics. (2018). *Catalyzing change in high school mathematics: Initiating critical conversations*. Author.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common core state standards for mathematics*. Author. Retrieved from http://corestandards.org/assets/CCSSI_Math%20Standards.pdf

National Research Council. (2001). *Adding it up: Helping children learn mathematics.* Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). Mathematics Learning Study Committee, Center for Education, Division of Behavioral and Social Sciences and Education. National Academy Press. http://dx.doi.org/10.17226/9822

Ni, Y., Zhou, D. R., Cai, J., Li, X., Li, Q., & Sun, I. X. (2018). Improving cognitive and affective learning outcomes of students through mathematics instructional tasks of high cognitive demand. *The Journal of Educational Research, 111*(6), 704–719. http://dx.doi.org/10.1080/00220671.2017.1402748

Nie, B., Freedman, T., Hwang, S., Wang, N., Moyer, J. C., & Cai, J. (2013). An investigation of teachers' intentions and reflections about using *Standards*-based and traditional textbooks in the classroom. *ZDM Mathematics Education, 45*, 699–711. http://dx.doi.org/10.1007/s11858-013-0493-7

Organisation for Economic Co-operation and Development. (2013). *Lessons from PISA 2012 for the United States: Strong performers and successful reformers in education.* Author.

Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and practice.* (4th ed.). Thousand Oaks, CA: Sage Publications.

Patton College of Education. (2019, November 4). *MODSPAR: Mathematical modeling and spatial reasoning.* https://www.ohio.edu/education/morton/modspar

Peressini, D., Borko, H., Romagnano, L., Knuth, E., & Willis, C. (2004). A conceptual framework for learning to teach secondary mathematics: A situative perspective. *Educational Studies in Mathematics, 56*(1), 67–96. http://dx.doi.org/10.1023/B:EDUC.0000028398.80108.87

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

Privitera, G. J., & Ahlgrim-Delzell, L. (2019). *Research methods for education.* Sage.

Remillard, J. T. (2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research*, *75*(2), 211–246. http://dx.doi.org/10.3102/00346543075002211

Resnick, L. B., & Hall, M. W. (1998). Learning organizations for sustainable education reform. *Daedalus, 127*, 89–118.

Richardson, V. (1996). The role of attitudes and beliefs in learning to teach. In J. Sikula (Ed.), *Handbook of research on teacher education* (pp. 102–119). Macmillan. http://dx.doi.org/10.12691/education-5-10-6

Rokeach, M. (1968). *Beliefs, attitudes, and values: A theory of organization and change.* Jossey-Bass. https://doi.org/10.1093/sw/14.4.115

Romagnano, L. (1994). *Wrestling with change: The dilemmas of teaching real mathematics*. Heinemann.

Rosenshine, B. (1980). How time is spent in elementary classrooms. In C. Denham & A. Lieberman (Eds.), *Time to learn* (pp. 107–126). National Institute of Education. https://doi.org/10.12691/education-5-10-3

Ross, J. A., McDougall, D., & Hogaboam-Gray, A. (2002). Research on reform in mathematics education, 1993–2000. *Alberta Journal of Educational Research, 48*, 122–138.

Rosswork, S. (1977). Goal setting: The effects on an academic task with varying magnitudes of incentive. *Journal of Educational Psychology, 69*, 710–715. http://dx.doi.org/10.1037/0022-0663.69.6.710

Rubin, H. J. & Rubin, I. S. (2012). *Qualitative interviewing: The art of hearing data* (3rd ed.). Sage.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4–14. http://dx.doi.org/10.3102/0013189X015002004

Silver, E. A., & Stein, M. K. (1996). The QUASAR project: The "revolution of the possible" in mathematics instructional reform in urban middle schools. *Urban Education, 30*(4), 476–521. http://dx.doi.org/10.1177/0042085996030004006

Smith, J. P. (1996). Efficacy and teaching mathematics by telling: A challenge for reform. *Journal for Research in Mathematics Education, 27*, 387–402. http://dx.doi.org/10.2307/749874

Smith, M. S., Bill, V., & Hughes, E. K. (2008). Thinking through a lesson: Successfully implementing high-level tasks. *Mathematics Teaching in the Middle School, 14*(3), 132–138. http://dx.doi.org/10.5951/mtms.14.3.0132

Smith, M. S., & Stein, M. K. (2018). *5 Practices for orchestrating productive mathematics discussions* (2nd ed.). National Council of Teachers of Mathematics.

Stein, M. K., Grover, B. W., & Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. *American Educational Research Journal*, *33*(2), 455–488. http://dx.doi.org/10.3102/00028312033002455

Stein, M. K., & Lane, S. (1996). Instructional tasks and the development of student capacity to think and reason: An analysis of the relationship between teaching and learning in a reform mathematics project. *Educational Research and Evaluation, 2*(1), 50–80. http://dx.doi.org/10.1080/1380361960020103

Stein, M., Remillard, J., & Smith, M. (2007). How curriclum influence student learning. In F. K. Lester Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 319–369). Information Age.

Stein, M. K., & Smith, M. S. (1998). Mathematical tasks as a framework for reflection: From research to practice. *Mathematics Teaching in the Middle School, 3*, 268–276. http://dx.doi.org/10.5951/MTMS.3.4.0268

Stein, M. K., Smith, M. S., Henningsen, M. A., & Silver, E. A. (2009). *Implementing standards-based mathematics instruction: A casebook for professional* *development* (2nd ed.). National Council of Teachers of Mathematics and Teachers College Press.

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. The Free Press.

Stigler, J. W., & Hiebert, J. (2004). Improving mathematics teaching. *Educational Leadership, 61*(5), 12–17.

Stylianides, A. J., & Delaney, S. (2018). Pre-service mathematics teachers' knowledge and beliefs. In G. J. Stylianides & K. Hino (Eds.), *Research advances in the mathematical education of pre-service elementary teachers: An international perspective* (pp. 219–228). Springer International Publishing AG. http://dx.doi.org/10.1007/978-3-319-68342-3_15

Tran, D., & Tarr, J. E. (2018). Examination of bivariate data tasks in US high school textbooks through the statistical investigation and cognitive demands frameworks. *International Journal of Science and Mathematics Education, 16*, 1581–1603. http://dx.doi.org/10.1007/s10763-017-9851-1

United States Census Bureau. (2021). *Fitting a line to data - Earnings and educational attainment.* https://www2.census.gov/programs-surveys/sis/activities/math/mm-14_teacher.pdf

Wideen, M., Mayer-Smith, J., & Moon, B. (1998). A critical analysis of the research on learning to teach: Making the case for an ecological perspective on inquiry. *Review of Educational Research, 68*(2), 130–178. http://dx.doi.org/10.3102/00346543068002130

Wilson, M. S., & Cooney, T. (2002). Mathematics teacher change and development, In G. C. Leder, E. Pehkonen, & G. Törner (Eds.), *Beliefs: A hidden variable in mathematics education* (pp. 127–147). Kluwer. http://dx.doi.org/10.1007/0-306-47958-3_8

Wilhelm, A. G., & Kim, S. (2015). Generalizing from observations of mathematics teachers' instructional practice using the Instructional Quality Assessment. *Journal for Research in Mathematics Education*, *46*(3), 270–279. http://dx.doi.org/10.5951/JRESEMATHEDUC.46.3.0270

Wolcott, H. (2005). *The art of fieldwork* (2nd ed.). AltaMira Press.

Zelkowski, J. (2009). Tackling low-level cognitive task teaching in secondary mathematics: An approach to shaping preservice teachers' preconceived beliefs. *Journal of Inquiry and Action in Education, 3*(1), 70–94.

Appendix A: The Instructional Quality Assessment Rubrics

Included with permission from Melissa Boston

Instructional Quality Assessment in Mathematics Classroom Observation Toolkit

IQA Mathematics Lesson Observation Rubrics

IQA Mathematics Rubrics for Student Work (or Assignment) Collections available as well.

IQA Mathematics Rubrics: Academic Rigor

RUBRIC AR1: Potential of the Task	
4	**The task has the potential to engage students in exploring and understanding the nature of mathematical concepts, procedures, and/or relationships, such as:** • Doing mathematics: using complex and non-algorithmic thinking (i.e., there is not a predictable, well-rehearsed approach or pathway explicitly suggested by the task, task instructions, or a worked-out example); OR • Procedures with connections: applying a broad general procedure that remains closely connected to mathematical concepts. **The task must explicitly prompt for evidence of students' reasoning and understanding.** For example, the task **MAY** require students to: • solve a genuine, challenging problem for which students' reasoning is evident in their work on the task; • develop an explanation for why formulas or procedures work; • identify patterns and form and justify generalizations based on these patterns; • make conjectures and support conclusions with mathematical evidence; • make explicit connections between representations, strategies, or mathematical concepts and procedures. • follow a prescribed procedure in order to explain/illustrate a mathematical concept, process, or relationship.
3	**The task has the potential to engage students in complex thinking or in creating meaning for mathematical concepts, procedures, and/or relationships. However, the task does not warrant a "4" because the task does not explicitly prompt for evidence of students' reasoning and understanding.** For example, students may be asked to: • engage in problem-solving, but the problem does not require much cognitive challenge (e.g., a problem that is easy to solve). • explore why formulas or procedures work, but not provide an explanation. • identify patterns, but are not pressed to explain generalizations or provide justification; • to make conjectures, but are not asked to provide mathematical evidence or explanations to support conclusions. • use multiple strategies or representations, but the task does not explicitly prompt students to develop connections between them. • follow a prescribed procedure to make sense of a mathematical concept, process, or relationship, but not to explain or illustrate the underlying mathematical ideas or relationships.
2	**The potential of the task is limited to engaging students in using a procedure that is either specifically called for or its use is evident based on prior instruction, experience, or placement of the task.** • There is little ambiguity about what needs to be done and how to do it. • The task does not require students to make connections to the concepts or meaning underlying the procedure being used. • **Focus of the task appears to be on producing correct answers rather than developing mathematical understanding (e.g., applying a specific problem solving strategy, practicing a computational algorithm).**

	OR There is evidence that the mathematical content of the task is at least 2 grade-levels below the grade of the students in the class.
1	**The potential of the task is limited to engaging students in memorizing or reproducing facts, rules, formulae, or definitions. The task does not require students to make connections to the concepts or meaning that underlie the facts, rules, formulae, or definitions being memorized or reproduced.**
0	**The task or lesson activity requires no mathematical activity.**

RUBRIC AR2: Implementation of the Task	
4	**Students engaged in exploring and understanding the nature of mathematical concepts, procedures, and/or relationships, such as:** • Doing mathematics: using complex and non-algorithmic thinking (i.e., there is not a predictable, well-rehearsed approach or pathway explicitly suggested by the task, task instructions, or a worked-out example); OR • Procedures with connections: applying a broad general procedure that remains closely connected to mathematical concepts. There is explicit evidence of students' reasoning and understanding. For example, students may have: • solved a genuine, challenging problem for which students' reasoning is evident in their work on the task; • developed an explanation for why formulas or procedures work; • identified patterns, formed and justified generalizations based on these patterns; • made conjectures and supported conclusions with mathematical evidence; • made explicit connections between representations, strategies, or mathematical concepts and procedures. • followed a prescribed procedure in order to explain/illustrate a mathematical concept, process, or relationship.
3	**Students engaged in complex thinking or in creating meaning for mathematical concepts, procedures, and/or relationships. However, the implementation does not warrant a "4" because there are no explicit explanations or written work to indicate students' reasoning and understanding. Students may have:** • engaged in problem-solving, but the problem did not require much cognitive challenge (e.g., the problem was easy to solve) or students' reasoning is not evident in their work on the task. • explored why formulas or procedures work, but did not provide explanations. • identified patterns but did not form or justify generalizations. • made conjectures but did not provide mathematical evidence or explanations to support conclusions • used multiple strategies or representations but connections between different strategies/representations were not explicitly evident; • followed a prescribed procedure to make sense of a mathematical concept, process, or relationship, but did not to explain or illustrate the underlying mathematical ideas or relationships.

2	**Students engaged in using a procedure that was either specifically called for or its use was evident based on prior instruction, experience, or placement of the task.** • There was little ambiguity about what needed to be done and how to do it. • Students did not make connections to the concepts or meaning underlying the procedure being used. • Implementation focused on producing correct answers rather than developing mathematical understanding (e.g., applying a specific problem solving strategy, practicing a computational algorithm). OR There is evidence that the mathematical content of the task is at least 2 grade-levels below the grade of the students in the class.
1	**Students engage in memorizing or reproducing facts, rules, formulae, or definitions. Students do not make connections to the concepts or meaning that underlie the facts, rules, formulae, or definitions being memorized or reproduced.**
0	The students did not engage in mathematical activity.

<table>
<tr><th colspan="2">RUBRIC AR3: Student Discussion Following Task</th></tr>
<tr><td>4</td><td><u>Students present their mathematical work and thinking</u> for solving a task and/or <u>engage in a discussion</u> (teacher-guided or student-led) of the important mathematical ideas in the task. During this discussion:
• students provide complete and thorough explanations of their strategy, idea, or procedure.
• students make connections to the underlying mathematical ideas (e.g., “I divided because we needed equal groups”).
• students provide reasoning and justification for their mathematical work and thinking.
OR
• students present and/or discuss more than one strategy or representation for solving the task, and a) provide explanations, comparisons, etc., of why/how the different strategies/representations were used to solve the task, and/or b) make explicit connections between strategies or representations;
• there is thorough presentation and discussion <u>across</u> strategies or representations</td></tr>
<tr><td>3</td><td><u>Students present their mathematical work and thinking</u> for solving a task and/or <u>engage in a discussion</u> (teacher-guided or student-led) of the important mathematical ideas in the task. During this discussion:
• students attempt to provide explanations of why their strategy, idea, or procedure is valid and/or students begin to make connections. The justifications, explanations and connections are conceptually-based (and on the right track), but are not complete and thorough (e.g., student responses often require extended press from the teacher, are incomplete, lack precision, or fall short of making explicit connections).
OR
• students present and/or discuss more than one strategy or representation for solving the task, and provide explanations of how the individual strategies/representations were used to solve the task but do not make connections between different strategies or representations.
• there are thorough presentation and/or discussion of individual strategies or representations, but there is not discussion, comparison, connections, etc., across strategies/representations.</td></tr>
<tr><td>2</td><td>Students show/describe/discuss procedural work for solving the task. During this discussion:
• connections are <u>not</u> made with mathematical concepts and the discussion focuses solely on procedures (e.g., the steps for a multiplication problem, finding an average, or solving an equation; what they did first, second, etc.), OR
• students make presentations of their work, and questioning or prompting from the teacher is for procedural explanations only, OR
• students show/discuss only one strategy/representation for solving the task, OR
• students present their work with no questioning or prompting from the teacher (to the presenters or to the class) to explain the mathematical work, make connections, etc. [Presentations with no discussion.]</td></tr>
<tr><td>1</td><td>• Students provide brief or one-word answers, fill in blanks, or IRE pattern (e.g., T: What is the answer to Question 5? S: 4.5 T: Correct!), OR
• Students’ responses are vague, unclear, or contain several misconceptions regarding the overall concept or procedure. [Student responses are incorrect or do not make sense mathematically.]</td></tr>
</table>

0	There was no mathematical discussion of the task: a) no discussion occurred following students' work on the task; or b) teacher's questions and/or student's responses are non-mathematical.

Rubric AR-Q: Rigor of Teacher's Questions	
4	**The teacher consistently asks** academically relevant questions that provide opportunities for students to elaborate and explain their mathematical work and thinking (probing, generating discussion), identify and describe the important mathematical ideas in the lesson, or make connections between ideas, representations, or strategies (exploring mathematical meanings and relationships).
3	**At least 3 times during the lesson, the teacher asks** academically relevant questions that provide opportunities for students to elaborate and explain their mathematical work and thinking (probing, generating discussion), identify and describe the important mathematical ideas in the lesson, or make connections between ideas, representations, or strategies (exploring mathematical meanings and relationships).
2	**There are one or more superficial, trivial, or formulaic efforts** to ask academically relevant questions probing, generating discussion, exploring mathematical meanings and relationships) (i.e., every student is asked the same question or set of questions) or to ask students to explain their reasoning; OR **Only one (1) effort** is made to ask an academically relevant question (e.g., one instance of a strong question, or the same strong question is asked multiple times)
1	The teacher asks procedural or factual questions that elicit mathematical facts or procedure or require brief, single word responses.
0	The teacher did not ask questions during the lesson, or the teacher's questions were not relevant to the mathematics in the lesson.
N/A	Reason:

Rubric AR-X: Mathematical Residue	
4	The discussion following students' work on the task surfaces the important mathematical ideas, concepts, or connections embedded in the task and serves to extend or solidify students' understanding of the main mathematical goals/ideas/concepts of the lesson. The discussion leaves behind important mathematical residue.
3	During the discussion following students' work on the task, the important mathematical ideas, concepts, or connections begin to surface, are wrestled with by students, but are not pursued in depth or have not materialized/solidified by the close of the lesson. The lesson is beginning to amount to something mathematically but the mathematics is only partially developed; perhaps due to time or student readiness.
2	During the discussion following students' work on the task, the important mathematical ideas, concepts, or connections in the task are explained or made explicit by the teacher primarily (i.e., the teacher is telling students what connections should have been made; students take notes or provide brief answers but do not make meaningful mathematical contributions to the discussion, students make superficial contributions that are taken over by the teacher). The discussion is mathematical, but does not address the concepts, ideas, or connections embedded in the task (random or not consistent with the mathematical goal) OR the discussion is about mathematics that is not relevant/important for the group of students.
1	Important mathematical ideas do not surface during the discussion following students' work on the task. The discussion is mathematical, but there is no apparent mathematical goal; the discussion does not focus on developing (or building up) students' understanding of the important mathematical ideas.
0	There was no discussion following the task. OR The discussion was about non-mathematical aspects of the task and did not leave behind mathematical residue.

IQA Mathematics Rubrics: Accountable Talk

Consider talk from the whole-group discussion only.

I. How effectively did the lesson-talk build Accountability to the Learning Community?

Participation in the Learning Community

Was there widespread participation in teacher-facilitated discussion?

Rubric 1: Participation	
4	Over 75% of the students participated throughout the discussion.
3	50-75% of the students participated in the discussion.
2	25-50% of the students participated in the discussion.
1	Less than 25% of the students participated in the discussion.
0	None of the students participated in the discussion.
N/A	Reason:

_______ Number of students in class

_______ Number of students who participated

Teacher's Linking Contributions: Does the teacher support students in connecting ideas and positions to build coherence in the discussion?

Rubric AT2: Teacher's Linking	
4	The teacher consistently (at least 3 times) explicitly connects (or provides opportunities for students to connect) speakers' contributions to each other and describes (or provides opportunities for students to describe) how ideas/positions shared during the discussion relate to each other.
3	At least twice during the lesson, the teacher explicitly connects (or provides opportunities for students to connect) speakers' contributions to each other and describes (or provides opportunities for students to describe) how ideas/positions relate to each other.
2	At one or more points during the discussion, the teacher links speakers' contributions to each other, but does not show how ideas/positions relate to each other (weak links -- e.g., local coherence; implicit building on ideas; noting that ideas/strategies are different but not describing how). OR teacher revoices or recaps only, but does not describe how ideas/positions relate to each other OR only one strong effort is made to connect speakers' contributions to each other (1 strong link).
1	Teacher does not make any effort to link or revoice speakers' contributions.
0	No class discussion OR Class discussion was not related to mathematics.
N/A	Reason:

Students' Linking Contributions: Do student's contributions link to and build on each other?

Rubric AT3: Students' Linking	
4	The students consistently explicitly connect their contributions to each other and describe how ideas/positions shared during the discussion relate to each other. (e.g. I agree with Jay because…")
3	At least twice during the lesson, students explicitly connect their contributions to each other and describe ideas/positions shared during the discussion relate to each other. (e.g. I agree with Jay because…")
2	At one or more points during the discussion, the students link students' contributions to each other, but do not describe how ideas/positions relate to each other. (e.g., e.g., local coherence; implicit building on ideas; "I disagree with Ana.") OR students make only one strong effort to connect their contributions with each other.

1	Students do not make any effort to link or revoice students' contributions.
0	No class discussion OR Class discussion was not related to mathematics.
N/A	Reason:

II. How effectively did the lesson-talk build <u>Accountability to Knowledge and Rigorous Thinking</u>?

Asking: Were students pressed to support their contributions with evidence and/or reasoning?

Rubric AT4: Teachers' Press	
4	The teacher consistently (almost always) asks students to provide evidence for their contributions (i.e., press for conceptual explanations) or to explain their reasoning. (There are few, if any instances of missed press, where the teacher needed to press and did not.)
3	Once or twice during the lesson the teacher asks students to provide evidence for their contributions (i.e., press for conceptual explanations) or to explain their reasoning. (The teacher sometimes presses for explanations, but there are instances of missed press.)
2	Most of the press is for computational or procedural explanations or memorized knowledge OR There are one or more superficial, trivial efforts, or formulaic efforts to ask students to provide evidence for their contributions or to explain their reasoning (i.e., asking everyone, "How did you get that?").
1	There are no efforts to ask students to provide evidence for their contributions AND there are no efforts to ask students to explain their thinking.
0	Class discussion was not related to mathematics OR No class discussion
N/A	Reason:

Providing: Did students support their contributions with evidence and/or reasoning? (This evidence must be appropriate to the content area—i.e., evidence from the text; citing an example, referring to prior classroom experience.)

Rubric AT5: Students' Providing	
4	Students consistently provide evidence for their claims, OR students explain their thinking using reasoning in ways appropriate to the discipline (i.e. conceptual explanations).
3	Once or twice during the lesson students provide evidence for their claims, OR students explain their thinking, using reasoning in ways appropriate to the discipline (i.e. conceptual explanations).

2	Students provide explanations that are computational, procedural or memorized knowledge, OR What little evidence or reasoning students provide is inaccurate, incomplete, or vague.
1	Speakers do not back up their claims, OR do not explain the reasoning behind their claims.
0	Class discussion was not related to mathematics OR No class discussion
N/A	Reason:

Appendix B: The Task as Set Up Rubrics

Included with the permission of Kara Jackson

The Task as Set Up rubrics are designed to accompany the *Instructional Quality Assessment* (IQA) middle-grades mathematics rubrics. To use the set-up rubrics, please contact Kara Jackson at karajack@uw.edu. To use the IQA rubrics, please contact Melissa Boston at bostonm@duq.edu.

Task as Set Up Rubrics

I. How effectively did the task-as-set-up phase of instruction* establish a shared understanding of the contextual features of the problem-solving scenario and what is to be mathematized such that the students are able to begin solving the task?
*****Includes prior to introduction of the task and/or in the context of introducing the task

***Contextual Features (Rubric 1): Establishing familiarity with the contextual features of the problem-solving scenario*:** To what extent were students supported to develop a shared understanding of the contextual features of the problem-solving scenario?

CONTEXTUAL FEATURES (RUBRIC 1): ESTABLISHING FAMILIARITY WITH THE CONTEXTUAL FEATURES OF THE PROBLEM-SOLVING SCENARIO	
4	Teacher elicits what students know about the problem-solving scenario. The teacher and/or students **consistently** make connections between ideas that are likely to support the establishment of a shared understanding of the contextual features of the problem-solving scenario (e.g., through accountable talk moves like marking, revoicing, linking, pressing; connecting to shared experiences; establishing a representation that students refer to). More than one student actively participates in this segment of instruction. Students respond in ways that demonstrate knowledge or understanding of the contextual features of the problem-solving scenario. **The teacher and/or students must consistently make connections between ideas (that could include the establishment of a representation) that support and/or build toward a shared understanding of the contextual features of the scenario.** *The difference between a 3 and a 4 is that the teacher and/or students **consistently connect** ideas that build toward a shared understanding of the contextual features of the problem-solving scenario.*
3	Teacher elicits what students know about the problem-solving scenario. The teacher and/or students **inconsistently** make connections between ideas that are likely to support the establishment of a shared understanding of the contextual features of the problem-solving scenario (e.g., through accountable talk moves like marking, revoicing, linking, pressing; connecting to shared experiences; establishing a representation that students refer to). More than one student actively participates in

	this segment of instruction. Students respond in ways that demonstrate knowledge or understanding of the contextual features of the problem-solving scenario. **The teacher and/or students inconsistently make connections between ideas (that could include the establishment of a representation) that support and/or build toward a shared understanding of the contextual features of the scenario.** *The difference between a 2 and a 3 is that the teacher and/or students connect ideas together (albeit inconsistently) to build toward a shared understanding of the contextual features of the problem-solving scenario.*
2	Teacher elicits what students know about the problem-solving scenario *but* the teacher and/or students **do not connect ideas** together in a way that would support students in establishing a shared understanding of the contextual features of the problem-solving scenario. (I.e., the teacher surfaces some initial ideas about the contextual features of the problem-solving scenario, but the ideas do not build to a greater or shared understanding of the contextual features. For example, the teacher may ask the same question to a number of students to gather information, "What is your favorite_____?" but does not support a shared understanding of the particular idea.) *The difference between a 1 and a 2 is that students must actively participate in the discussion to warrant a 2.*
1	Teacher makes at least a brief mention of the problem-solving scenario that is central to completion of the task. The teacher is the *only* person providing information about the contextual features of the problem-solving scenario. The teacher may ask questions that require yes/no responses. Students do not actively participate in this chunk of instruction. *Examples:* • Superficial attempt (*Does anyone have a cell phone? Show a picture of a cell phone*.)
0	There is no attempt to discuss the contextual features of the problem-solving scenario.
N/A	The task as provided does not have a problem-solving scenario (e.g., it is a set of naked number problems).
NS	(No score). There is <u>no whole class discussion of the task prior to students starting the task</u> (e.g., teacher hands out the task and tells the students to start the task; teacher hands out the task and has students discuss it in groups prior to working on the task, but there is no whole class discussion prior to students starting the task). **Only assign NS for Contextual Features if the task has a problem-solving scenario* and *there is no whole class discussion of the task prior to students starting the task.*

MATHEMATICAL RELATIONSHIPS (Rubric 2): Developing understandings of key mathematical relationships as they are represented in the task: To what extent have students been given opportunities to develop understandings of key mathematical relationships (e.g., key mathematical ideas, relationships and/or quantities) as represented in the task?

MATHEMATICAL RELATIONSHIPS (RUBRIC 2): DEVELOPING UNDERSTANDINGS OF WHAT IS TO BE MATHEMATIZED IN THE TASK	
4	Teacher elicits the ideas students have developed *and* the teacher supports students in establishing a shared understanding of key mathematical ideas, relationships, and/or quantities (e.g., through marking, revoicing, linking, pressing) as represented in the task. Students actively participate in this segment of instruction. Students respond in ways that demonstrate understanding of how key mathematical ideas, relationships, and/or quantities are represented in the task. The students' responses connect to and build on each other. **In addition to what qualifies as a 3, the teacher and/or students must make at least one strong accountable talk move (e.g., teacher and/or student identifies connections between ideas and *how* the ideas are related, teacher presses on student's understandings of the mathematical ideas).** *The difference between a 3 and a 4 is the presence of* ***one strong accountable talk move*** *on the part of the teacher or students.*
3	Teacher elicits information about the ideas students have developed. The teacher and/or students **consistently use accountable talk moves (other than, or in addition to, repeating students' contributions)** to support students in establishing a shared understanding of how key mathematical ideas, relationships, and/or quantities are represented in the task. **Alternatively, the students (with or without the involvement of the teacher) jointly establish a representation that supports a shared understanding of relevant mathematical relationships**. Students actively participate in this segment of instruction. Students respond in ways that demonstrate understanding of how key mathematical ideas, relationships, and/or quantities are represented in the task. **The teacher and or/students consistently use accountable talk moves (other than, or in addition to, repeating students' contributions).** **OR** **More than one student is involved in the establishment of a representation that supports a shared understanding of mathematical relationships relevant to what is to be mathematized. Student talk must be aimed at developing a conceptual understanding of relevant mathematical ideas. (*More than one student must be involved in conceptual talk related to the representation.)** *The difference between a 2 and a 3 is that the teacher and/or students consistently use accountable talk moves (other than, or in addition to, repeating students'*

	contributions) ***or*** *students are involved in the* ***joint establishment*** *of a representation to develop a shared understanding of the mathematical relationships.*
2	Teacher elicits information about the ideas students have developed *but* the students (with or without the involvement of the teacher) do not **jointly establish** a representation nor are there consistent accountable talk moves that would support students in establishing a shared understanding of how key mathematical ideas, relationships, and/or quantities are represented in the task. **At best, there is inconsistent use of accountable talk moves (or consistent repeating of students' contributions).** *The difference between a 1 and a 2 is that students must actively participate in the discussion to warrant a 2.*
1	Teacher makes at least a brief attempt to provide or suggest how key mathematical ideas, relationships, and/or quantities are represented in the task. The teacher is the *only* person suggesting or providing ideas regarding how mathematical ideas, relationships, and/or quantities are represented in the task. The teacher may ask questions that require brief or one-word answers. Students do *not* actively participate in this chunk of instruction. At most, students provide yes/no responses or nod heads.
0	There is no attempt to discuss key mathematical ideas, relationships, and/or quantities.
N/A	The task is not mathematical in nature.
NS	(No score). There is <u>no whole class discussion of the task prior to students starting the task</u> (e.g., teacher hands out the task and tells the students to start the task; teacher hands out the task and has students discuss it in groups prior to working on the task, but there is no whole class discussion prior to students starting the task).

Appendix C: Pre-Observation Interview Protocol

Teacher Information

Teacher's name:

Teacher's education and degrees earned:

Teachers' years of experience (and at this school):

Teacher's course load:

Teacher's role in MMR course (how many years piloting?):

Lesson Information

Lesson title:

Course name:

Grade level(s):

Main task title:

Lesson content addressed:

Task Selected from Source Materials:

- What things do you look for when selecting tasks or activities to use with your students?
    - Why are these things important to you?
    - How often do you use tasks with your students? What other teaching approaches do you use? How do you determine which to use and when?
- What led you to choose to use this task with your students? *Would you use it even if it was not a required part of the MMR course materials?
    - How does this task help students accomplish their learning goals?

  - How does this task address course or unit objectives?
  - Why do you believe that the math content (or other content) of this task is important for students to learn?
  - How does this task address math concepts? (how successful do you think it will be?)
  - How does this task address student engagement? (how successful do you think it will be?)
  - What features make this task desirable for you? What task features are undesirable?
- Where was this task taken or adapted from?
  - Where did you first learn of this resource?
  - How long have you been using this resource?
    - Why do you continue to use it?
    - How comfortable are you with using or adapting it?
  - How often do you use this resource?
  - Do you think this resource is effective? Why or why not?
  - If you created the task yourself, what was your motivation for doing so? How did you determine what to include and not include? Did you borrow some ideas from other resources? Explain.
- What math representations (e.g., graphs, numbers, equations) are possible to use with this task? Which might you emphasize (or not emphasize)? Why?

- Classify this version of the task using the Task Analysis Guide and explain why your rating is appropriate.
    - **If *Memorization*:** what information needs to be memorized to succeed on this task?
    - **If *Procedures without Connections*:** what procedure(s) are involved in this task? What other procedures could be used?
    - **If *Procedures with Connections*:** what procedure(s) are involved in this task? What other procedures could be used? How are the procedures connected to mathematical concepts or relationships?
    - **If *Doing Mathematics*:** how does this task exceed the level of *procedures with connections*? (e.g, how does it involve complex, nonalgorithmic thinking?)

Task Planning

- Can you tell me about how you plan for class every day?
    - What makes you plan this way?
    - How far in advance do you plan your lessons?
    - How has your planning changed as you have gained teaching experience?
    - Can you tell me about your students who will be doing this task?
        - How do you think they will do with it? Why do you think so?
- How have you modified the original task for use in your own classroom?
    - What changes did you make during your lesson planning?

    - Why did you make these changes?
    - How would your proposed changes to the task influence the potential for student learning?
        - What else might they learn?
        - Why do you think so?
        - How else might these changes influence students' engagement with the task?
- Have you taught this content before?
    - If so, have you done it with the same (or a similar) lesson plan? Describe what is different (or similar) about this task compared to what you have done in the past with similar content.
    - If you modified the lesson plan from before, how and why?
    - How comfortable are you teaching this content? This activity or task?
- What objectives would you have for students who worked on this task?
    - How would you know that students have met these objectives?
    - Why are these objectives important?
    - What are your expectations for student work when solving this task? How do you plan on making these expectations clear to students?
- What difficulties or problems would you anticipate students might have when working on this task?
    - Why do you think these difficulties or problems might occur?
    - How might you handle them? Why would you handle them this way?

- How might your students' prior knowledge help them to succeed in solving this task?
    - What new information might students need to complete the task?
- How do you plan on setting this task up for your students?
    - For example, what additional instructions might you give them? What tools and resources will you make available (or not available)? Why?
    - Will you have students work in small groups? Individuals? Pairs? Why?
    - What background information will you provide to students?
    - What questions will you ask?
    - What math representations (e.g., graphs, numbers, equations) will you make available for students? What representations will you emphasize (or not emphasize)? Why?
- Classify this version of the task using the Task Analysis Guide and explain why your rating is appropriate.
    - **If *Memorization*:** what information needs to be memorized to succeed on this task?
    - **If *Procedures without Connections*:** what procedure(s) are involved in this task? What other procedures could be used?
    - **If *Procedures with Connections*:** what procedure(s) are involved in this task? What other procedures could be used? How are the procedures connected to mathematical concepts or relationships?

- **If *Doing Mathematics*:** how does this task exceed the level of *procedures with connections*? (e.g, how does it involve complex, nonalgorithmic thinking?)
- How does this classification compare with your classification for the original task?
- Why do you think your rating for this task is the same or different than it was before? Explain.

Appendix D: Post-Observation Interview Protocol

Teacher Information

Teacher's name:

Teacher's education and degrees earned:

Teachers' years of experience (and at this school):

Teacher's course load:

Teacher's role in MMR course (how many years piloting?)

Lesson Information

Lesson title:

Course name:

Grade level(s):

Main task title:

Lesson content addressed:

Task Setup:

- Can you tell me about how you set this task up for your students?
    - Did you set the task up in the same way that you had planned to? How was the setup task the same or different than you had planned it to be? Why do you think so?
    - How was the setup of this task similar or different than other tasks you have launched? Do you typically give the same amount of instruction or guidance? Why or why not?
- How did your students respond to the setup of the task?

  - What were some of the questions they asked? Why do you think they asked those questions?
  - Did students engage with the task readily? Did they require additional clarification or support? Why do you think so?
- Did students have access to the same tools and resources as you anticipated they would? Why or why not?
- What questions did you ask students when setting up the task?
  - How did students respond?
  - Did you anticipate them to respond in this way? What did you do when students said or did something unexpected?
- What math representations did you make available for students (e.g., graphs, numbers, equations)? What representations did you emphasize? Why?
- Classify this version of the task using the Task Analysis Guide and explain why your rating is appropriate.
  - **If *Memorization*:** what information needs to be memorized to succeed on this task?
  - **If *Procedures without Connections*:** what procedure(s) are involved in this task? What other procedures could be used?
  - **If *Procedures with Connections*:** what procedure(s) are involved in this task? What other procedures could be used? How are the procedures connected to mathematical concepts or relationships?

    - **If *Doing Mathematics*:** how does this task exceed the level of *procedures with connections*? (e.g, how does it involve complex, nonalgorithmic thinking?)
    - How does this classification compare with the task as planned?
    - Why do you think your rating for this task is the same or different than it was before? Explain.

Task Implementation:

- Walk me through how this task played out in the classroom.
    - How did students' engagement with the task compare with what you had anticipated? How prepared were you to interact with students during the lesson?
    - How did students' engagement with the task compare to their engagement with other similar tasks you have used before? What was the same or different? Why do you think so?
    - Did students meet your learning goals and objectives for the task? How do you know?
    - How successful were students at identifying and understanding the math concepts embedded in the task? How do you know?
    - Describe the student work and thinking that resulted from their engagement with the task. Did their work and thinking meet your expectations? What did they do well? What could they have done better? Why do you think so?

    - Did students make use of the tools and resources you made available to them? How effective were they at using such resources? Were there any resources that they did not use? Why do you think they chose some resources over others?
    - Did anything about the lesson or task stand out to you? Why or how?
- Explain why you had students work in groups (or individually) and how successful you feel they were with their work.
    - Did students cooperate and engage with each other productively? What did you do to support their teamwork and engagement? How helpful was it?
- Can you recall an instance where a student(s) struggled with the task? Describe it.
    - How did you respond in this instance?
    - Why did you respond in this way?
    - Were the student(s) able to overcome the obstacle? How do you know?
    - How well do you think you supported productive struggle with the task? What evidence of productive struggle was evident during the lesson?
- What challenges did you face when implementing this lesson?
    - Why do you think you faced these challenges?
    - How did you address them?
- If you facilitated a whole-class discussion, how effective do you think it was?
    - Why do you think so?
    - How did you determine which student(s) should present and when?

    - What contributions did you make to the discussion and how did you know when to step in (or step out)?
- What math representations were used by you or the students (e.g., graphs, numbers, equations)?
    - Why do you think those were used?
    - Were there representations you expected students to use but were not used?
- Classify this version of the task using the Task Analysis Guide and explain why your rating is appropriate.
    - **If *Memorization*:** what information needs to be memorized to succeed on this task?
    - **If *Procedures without Connections*:** what procedure(s) are involved in this task? What other procedures could be used?
    - **If *Procedures with Connections*:** what procedure(s) are involved in this task? What other procedures could be used? How are the procedures connected to mathematical concepts or relationships?
    - **If *Doing Mathematics*:** how does this task exceed the level of *procedures with connections*? (e.g, how does it involve complex, nonalgorithmic thinking?)
    - How does this classification compare with the task as it was set up?
    - Why do you think your rating for this task is the same or different than it was before? Explain.

- What do you believe makes a task, activity, or lesson that you have taught "successful?"
- How do you know that your students are learning and making sense of the math throughout a lesson or task?
    - After the lesson or task is finished?
    - In the case of this particular task?
- If you were to implement this task again, what things would you change?
    - Why would you change these things?
    - How do you think these changes would help to maintain or enhance the cognitive demand of the task?

Appendix E: Ohio University Research Consent Form

Title of Research:

High School Mathematics Teachers' Perspectives on Selecting, Planning, Setting Up, and Implementing High Cognitive Demand Instructional Tasks

Researchers: Otto Shaw (Primary Investigator), Harman P. Aryal (Co-Investigator), and Gregory D. Foley (Adviser)

IRB number: 20-E-248

You are being asked by an Ohio University researcher to participate in research. For you to be able to decide whether you want to participate in this project, you should understand what the project is about, as well as the possible risks and benefits, in order to make an informed decision. This process is known as informed consent. This form describes the purpose, procedures, possible benefits, and risks of the research project. It also explains how your personal information will be used and protected. Once you have read this form and your questions about the study are answered, you will be asked to participate in this study. You should receive a copy of this document to take with you.

Summary of the Study

The primary investigator (PI) aims to study high school mathematics teachers' perspectives as they select, plan, set up, and implement high cognitive demand instructional tasks (e.g., problems or activities that support high-level student thinking). The PI also seeks to learn how such teachers reflect on their instructional decisions throughout the process. The PI will interview participating teachers, during which you can explain how and why you made instructional decisions as you selected and planned tasks to use with your students. Your teaching will also be observed on two consecutive days by the PI and the co-investigator (co-I) and they will ask you to submit 6–12 samples of student work from the observed lessons. Finally, the PI will interview you again after observing your instruction to explore how you set up and implemented the tasks.

Explanation of the Study

This study is being done to investigate high school mathematics teachers' perspectives as they use instructional tasks, with the purpose of providing insights that might support teacher professional development and teacher education programs. Your responses to interview questions and use of instructional tasks may also help to assist other teachers who use instructional tasks with their students. If you agree to participate, you will be asked to do two interviews, each lasting no longer than 1 hr, focusing on your instructional decisions as you select, plan, set up, and implement instructional tasks. You will also be asked to allow the PI and co-I to observe you teach on two consecutive days and collect student work samples from each lesson observed. You should not participate in this study if you are not a licensed high school mathematics teacher in the state of Ohio or if you are not teaching high school mathematics in the state of Ohio at this time. All contact and data collection procedures will occur remotely to minimize health risks due to COVID-19.

Risks and Discomforts

No risks or discomforts are anticipated.

Benefits

You may not benefit personally by participating in this study. However, you will be encouraged to reflect on your instructional decisions and may benefit from insights you obtain through self-reflection. This study might offer ways to support teacher professional development workshops, teacher preparation programs, and other mathematics teachers attempting to use high-level instructional tasks.

Confidentiality and Records

Your study information will be kept confidential by the assignment of pseudonyms in any written report based on this research. Your interview transcript(s) will be shared with you to verify transcription accuracy and as an invitation for you to add to, remove from, or modify them and ensure that they accurately reflect your thoughts and opinions. Names and identifiers will be removed from all student work collected and analyzed as part of this study. This will be done either by research participants prior to sending the samples to the PI or the PI will remove them immediately when he receives them.

Interviews will be audio-recorded—or video recorded if done by computer—and transcribed into text so that the PI can refer back to them as needed. Teaching observations will be recorded through written field notes and will not be audio or visually recorded. Recordings will be stored on the laptop and flash-drive of the PI and will not be accessible to anyone else. All audio, video, and digital files containing interview data and student work will be destroyed in May, 2022.

Additionally, although every effort will be made to keep your study-related information confidential, there may be circumstances where this information must be shared with:

Federal agencies, for example the Office of Human Research Protections, whose responsibility is to protect human subjects in research;

Representatives of Ohio University (OU), including the Institutional Review Board, a committee that oversees the research at OU.

Compensation

No compensation will be provided.

Future Use Statement

Data without identifiers may be used for future research studies or distributed to another investigator for future research studies without additional informed consent from you or your legally authorized representative.

Contact Information

If you have any questions regarding this study, please contact the investigators or adviser:

Primary Investigator	**Co-Investigator**	**Adviser**
Otto Shaw	Harman P. Aryal	Dr. Gregory D. Foley
os005910@ohio.edu	ha333416@ohio.edu	foleyg@ohio.edu
740.583.4703	740.707.4097	740.593.4430

If you have any questions regarding your rights as a research participant, please contact Dr. Chris Hayhow, Director of Research Compliance, Ohio University, (740)593-0664 or hayhow@ohio.edu.

By agreeing to participate in this study, you are agreeing that:

you have read this consent form (or it has been read to you) and have been given the opportunity to ask questions and have them answered;

you have been informed of potential risks and they have been explained to your satisfaction;

you understand Ohio University has no funds set aside for any injuries you might receive as a result of participating in this study;

you are 18 years of age or older;

your participation in this research is completely voluntary;

you may leave the study at any time; if you decide to stop participating in the study, there will be no penalty to you and you will not lose any benefits to which you are otherwise entitled.

Version Date: 08/13/2020

Appendix F: Email Text Requesting Letter of Support from Principals

Dear (principal's name) and (teacher's name),

I hope that the school year is off to a good start.

I am an Ohio University student researcher conducting research entitled "High School Mathematics Teachers' Perspectives on Selecting, Planning, Setting Up, and Implementing High Cognitive Demand Instructional Tasks" (IRB no. 20-E-248) and wish to invite (teacher's name) to participate in the study.

Their participation would require two 1-hour interviews and would require me to observe their teaching on two consecutive days **(all done remotely)**. I would also collect one or two problems or activities that they use with their students during those two days and some samples of student work that go with them. All names and identifiers will be removed from artifacts that I collect to preserve teacher and student anonymity. **Additionally, all research and data collection procedures will occur remotely to avoid health risks due to COVID-19.** Would you please discuss with (teacher's name) if they might be interested in participating in this study?

Ohio University requires that all research must first be accepted through their Institutional Review Board (IRB); doing so requires a letter of support from a principal or higher before doing research involving teachers and students. If (teacher's name) might be interested and you are willing, this letter would need to specifically include a letter written on letterhead of your school or district, which—

- Is dated,
- Has your written signature,
- Clearly states your title as principal, and
- Includes a statement of your agreement allowing me to conduct research within your high school.

You may use the attached document as a template or may draft another letter meeting the above criteria.

If I receive your approval and the associated letter, I will send a similar email message to (teacher's name) asking for their consent to participate in this study. Thank you for your consideration. I look forward to hearing back from you.

Best,

Otto Shaw

Primary Investigator

Appendix G: Template Letter of Support to Conduct Research

To the Ohio University Institutional Review Board (IRB):

I am familiar with Otto Shaw's research project entitled "High School Mathematics Teachers' Perspectives on Selecting, Planning, Setting Up, and Implementing High Cognitive Demand Instructional Tasks". I understand teachers' involvement to be interviewed and observed teaching and I understand that students will be observed, and some samples of their in-class work will be collected. I understand that, because this is educational research, parental consent will not be required when observing students in class and collecting samples of their work, as per the Ohio University IRB guidelines.

I understand that this research will be carried out following sound ethical principles and that participant involvement in this research study is strictly voluntary and provides confidentiality of research data, as described in the protocol.

Therefore, as a representative of [agency/institution name], I agree that Otto Shaw's research project may be conducted at our agency/institution.

Sincerely,

[name and title of agency/institutional authority]

[signature]

[date]

Appendix H: Email Text Requesting Participation in Research (Ohio MMR)

Dear …,

I hope that the school year is off to a good start.

I am an Ohio University student researcher conducting research entitled "High School Mathematics Teachers' Perspectives on Selecting, Planning, Setting Up, and Implementing High Cognitive Demand Instructional Tasks" and wish to know if you might be interested in participating in the study (IRB no. 20-E-248). I wish to learn from the perspectives of teachers who are involved in the Ohio Mathematics Modeling and Reasoning (Ohio MMR) pilot course this year and how you are using mathematical tasks in your instruction.

Your participation would require two 1-hour interviews and would require me to observe you teach the Ohio MMR course on two consecutive days. I would also collect one or two tasks that you use with your students during the two observation days and some samples of student work that go with them. **All research and data collection procedures will be done remotely to prevent health risks due to COVID-19.**

Please let me know if you would like to be involved in this research officially and we can discuss your participation in more depth. Please see the attached Research Consent Form and let me know if you have any questions or concerns. Thank you for your consideration. I look forward to hearing back from you.

Best,

Otto Shaw

Primary Investigator

Appendix I: Email Text Requesting Participation in Research (ATC)

Dear …,

I hope that the school year is off to a good start.

I am an Ohio University student researcher conducting research entitled “High School Mathematics Teachers’ Perspectives on Selecting, Planning, Setting Up, and Implementing High Cognitive Demand Instructional Tasks” and wish to know if you might be interested in participating in the study (IRB no. 20-E-248). I would like to learn from the perspectives of teachers who had been involved in Advanced Teacher Capacity (ATC) programs at Ohio University, namely, Modspar, QUANT, or both. In particular, I would like to learn about how you are using mathematical tasks in your instruction.

Your participation would require two 1-hour interviews and would require me to observe your teaching on two consecutive days. I would also collect one or two tasks that you use with your students during the two observation days and some samples of student work that go with them. **All research and data collection procedures will be done remotely to prevent health risks due to COVID-19.**

Please let me know if you would like to be involved in this research and we can discuss your participation in more depth. Thank you for your consideration. I look forward to hearing back from you.

Best,

Otto Shaw

Primary Investigator