The role of content-focused coaching in fostering ambitious mathematics teaching practices in elementary classrooms

by

Corinne Marko Murawski

Bachelor of Science degree, The Pennsylvania State University, 1989

Master of Education degree, Duquesne University, 1999

Submitted to the Graduate Faculty of the

School of Education in partial fulfillment

of the requirements for the degree of

Doctor of Education

University of Pittsburgh

2019

UNIVERSITY OF PITTSBURGH

SCHOOL OF EDUCATION

This dissertation was presented

by

Corinne Marko Murawski

It was defended on

December 6, 2019

and approved by

Dr. Margaret S. Smith, Professor Emeritus, Instruction and Learning

Dr. Melissa Boston, Professor, Duquesne University

Dr. Keith Trahan, Professor, Collaborative for Evaluation and Assessment Capacity

Dissertation Director: Dr. Ellen Ansell, Professor, Instruction and Learning

The role of content-focused coaching in fostering ambitious mathematics teaching practices in elementary classrooms

Corinne Marko Murawski, EdD

University of Pittsburgh, 2019

Evidence exists that ambitious teaching in mathematics classrooms makes a difference for student learning, regardless of grade level, yet many teachers of mathematics do not employ such techniques. While there are multiple possible explanations for this, no *one* explanation has been proven. Additionally, some studies *have* shown success in helping teachers change pedagogical practices to implement more ambitious practices. This study proposes that teachers need more in situ professional learning in conjunction with outside-the-classroom professional development to catalyze a change in practice. One form of such in situ learning is content-focused coaching. This study compared teachers' practices to attempt to show the increased effects of content-focused coaching plus outside-the-classroom professional development in contrast to only the outside-the-classroom professional development component.

Findings clearly showed that coached teachers had significantly more opportunities to learn about ambitious teaching practices than comparison teachers. However, results were not as clear when the use of such practices was assessed. While coached teachers significantly improved scores on Academic Rigor (AR) rubrics from the Instructional Quality Assessment (IQA) and had better scores than their counterparts in the comparison group, teachers scores on the composite IQA did not significantly improve and were not better than the uncoached teachers' scores. Qualitatively, coached teachers' experiences with the effective Mathematics Teaching Practices during coaching were different than the experiences of the uncoached teachers.

Findings from this study demonstrates that coaching matters. The content and the quality of what happens during coaching around the effective teaching practices for mathematics impacts teachers' classroom practice. In addition, this study shows that pairing coaching with outside-the-classroom professional development that also exposes teachers to ambitious teaching practices helps teachers better implement the practices. This is particularly true when the pairing is purposeful. In other words, when the same ambitious practice(s) is the focus of coaching and concurrently the focus of outside-the-classroom professional development, teachers more readily implement the practice(s) in their classroom. Purposefully integrating coaching with the content of teachers' curriculum, and purposefully integrating coaching with teachers' current position along a trajectory for learning about ambitious teaching practices also helps teachers more readily implement ambitious instructional methods.

Table of Contents

List of Tables

List of Figures

1.0 Introduction

Over the last decades, advances in fields as seemingly disparate as neuroscience, anthropology, and psychology have helped us determine how people, including students, best learn. As *How People Learn* tells us, these advances have "important implications for education. …[A] new theory of learning is coming into focus that leads to very different approaches to the design of curriculum, teaching, and assessment than those often found in schools today" (Bransford, Brown, Cocking, & National Research Council, 1999, p. 3). Darling-Hammond and Bransford echo this sentiment by stating, "great strides have been made in our understanding of learning and the teaching practices that support it" (Darling-Hammond & Bransford, 2005).

Advances in learning theory include, but are not limited to, findings from research in mathematics education. *Adding It Up* succinctly summarizes some of the pertinent findings by stating, "The effectiveness of mathematics teaching and learning is a function of teachers' knowledge and use of mathematical content, of teachers' attention to and work with students, and of students' engagement in and use of mathematical tasks" (National Research Council, 2005, p. 9). However, while the field now has a more refined theory of how students learn mathematics (Findell, 2002; Lester Jr., 2007) and understands that teachers are a dominant contributing factor to student academic gain (Hiebert & Grouws, 2007; Wright, Horn, & Sanders, 1997), teachers of K-12 mathematics often do not employ practices that align with how students best learn (Hiebert et al., 2005)."Traditional models of instruction still dominate the educational landscape" (Staples, 2007, p. 161). The mathematics learning of another generation is at stake. Many adults unabashedly admit, "I was never good at math." In the globally competitive world of the twenty-first century, it is no longer an option for students to head towards a similar fate. To reverse course, and create a generation that understands and can use mathematics, teachers must change the way they teach mathematics in school.

Professional development for teachers of mathematics is plentiful, and some of it aligns with research and publications on teacher professional learning (Borasi & Fonzi, 2002; Desimone, 2009; Loucks-Horsley, Stiles, Mundry, Love, & Hewson, 2012; M. S. Smith, 2001) and student learning (Boston & Smith, 2011; Franke, Carpenter, Levi, & Fennema, 2001; Lampert et al., 2013). Despite this, a large proportion of mathematics teachers are still using traditional lecture or recitation methods (Horizon Research Inc., 2013). For some teachers, traditional methods persist even after having participated in professional development around best practices for teaching mathematics (Removcik, 2014; Wang & Romero, 2013). Ignoring best practices leads to the area of concern for the proposed research study. To introduce the problem of practice for this dissertation, the chapter first turns to a short, historical examination of the development of ambitious teaching practices. Following that, the chapter turns to examining the context of the problem that will be addressed in this dissertation including possible reasons the problem exists and potential solutions.

1.1 Ambitious teaching practices

Beginning before the publication of *A Nation at Risk* (The National Commission on Excellence in Education, 1983), there has been an increasing body of evidence that traditional methodologies for teaching mathematics are not effective with all students. In the 1980s, research aimed at improving instruction in mathematics came to the forefront with the publication of the NCTM's *Curriculum and Evaluation Standards for School Mathematics* (National Council of

Teachers of Mathematics [NCTM], 1989) which made the case for changes in instructional practice, in part, by stating,

> All industrialized countries have experienced a shift from an industrial to an information society, a shift that has transformed both the aspects of mathematics that need to be transmitted to students and the concepts and procedures they must master if they are to be self-fulfilled, productive citizens in the next century. (p. 3)

Prior to publication of the *Standards*, researchers at institutions across the country (e.g., Stanford, University of Pittsburgh) had already begun working with teachers to change the face of mathematics teaching (e.g., Cohen, 1990; Shulman, 1987; Stein & Wang, 1988). While the findings from a number of studies encouraged similar pedagogies, different labels were used over the years to describe such teaching. Terms such as “reform-oriented” (Stein, Grover, & Henningsen, 1996), based on the fact that mathematics teaching was reforming, and “standards-based” (Resnick, Stein, & Coon, 2008), due to the fact that these practices were encouraged by the NCTM Standards (NCTM, 1989, 1991, 1995, 2000), were used to describe practices involving similar methodologies.

Hiebert, et al. (1997) defined reform-oriented instruction as including: (1) the use of problematic tasks as chosen by the teacher; (2) the establishment of a culture of collaboration for learning; (3) the use of mathematical tools to support learning; and (4) mathematical discourse. Standards-based instruction, as defined in the *Journal for Research in Mathematics Education* includes an “emphasis on student-centered instruction that engages students in exploration of mathematical facts and principles through collaborative work on authentic problems” (Huntley, Rasmussen, Villarubi, Sangtong, & Fey, 2000, p. 329) wherein students construct meaning for the mathematical concepts and procedures they are investigating and engage in meaningful problem-solving activities…facilitated by teachers who elicit, support, and extend children's mathematical thinking (Fraivillig, Murphy, & Fuson, 1999); promote discussions (e.g., Schifter & O'Brien, 1997); use meaningful representations of mathematical concepts (Fuson, Smith, & Cicero, 1997; Fuson, Wearne, et al., 1997); and encourage use of alternative solution methods (Carpenter & Fennema, 1991; Hiebert & Carpenter, 1992). (Fuson, Carroll, & Drueck, 2000, p. 277). As one can see, the teaching pedagogies for reform-oriented mathematics instruction and standards-based mathematics instruction are parallel.

As the body of educational research increased and the impetus to have a common set of content-related expectations reached critical mass, a new set of standards, the Common Core State Standards (CCSS) (National Governors Association Center for Best Practices and Council of Chief State School Officers [NGA and CCSSO], 2010), came into existence. Within the CCSS, the exposition and specification of habits of mind related to student learning, some common across content areas and some not (e.g., Standards for Mathematical Practice, Capacities for English Language Arts), remained critical guideposts for instruction. The teaching methods for allowing students to engage in the Standards for Mathematical Practice are closely related to what was known as "reform-oriented" or "standards-based" teaching. These teaching methodologies are now characterized as "ambitious instruction."

There are many groups with ideas about ambitious teaching (e.g., Ball & Forzani, 2009; Windschitl, Thompson, Braaten, & Stroupe, 2012 etc.,). As one looks across the work of these groups, the teaching practices associated with ambitious teaching become numerous. For this inquiry, the author employs the following definition of ambitious teaching offered by Lampert, Boerst, and Graziani (2011). " 'Ambitious teaching' is teaching that aims to teach all kinds of students to not only know academic subjects but also to be able to use what they know in working on authentic problems in academic domains" (p. 1361). Furthermore, this study will narrow the field of ambitious teaching practices by using the eight effective Mathematics Teaching Practices (NCTM, 2014) as a means of concretizing how ambitious teaching in mathematics is enacted in the classroom. As stated by Smith, Boston, and Huinker (2017) in the preface of the book series *Taking Action: Implementing Effective Mathematics Teaching Practices,* "Decades of empirical research in mathematics classrooms support these teaching practices" (p. v). Using the eight effective Mathematics Teaching Practices as indicative of ambitious instruction will serve to limit the number of methodologies teachers must consider while improving their teaching and making it more ambitious[1]. Throughout this document, the term "ambitious teaching" is used in reference to studies and ideas that may precede the use of such terminology, but which refer to teaching mathematics in ways that are synonymous with the current meaning of ambitious teaching.

1.2 Context of the problem

Ambitious teaching practices in mathematics (Lampert, Beasley, Ghousseini, Kazemi, & Franke, 2010; Lampert et al., 2011) result in higher student achievement. Boaler and Staples, (2008) provided evidence for this claim. Their work in high schools showed that students in classrooms where teachers employ ambitious mathematics teaching practices have "higher overall achievement on a number of measures" (p. 608). While Boaler and Staples made a persuasive argument for ambitious instruction, they are not the only ones to provide research, information,

[1] Over decades, the author-researchers whose work inspired this study have been instrumental in developing, defining, and elaborating the eight effective Mathematics Teaching Practices within the volumes of the *Taking Action* series (Boston, Dillon, Smith, & Miller, 2017; Huinker & Bill, 2017; Smith, Steele, & Raith, 2017).

and data to support it. Boaler and Staples corroborated Stein, Grover, and Henningsen's (1996) findings from research connected to middle school mathematics in the Quantitative Understanding: Amplifying Student Achievement and Reasoning (QUASAR) project (Silver & Stein, 1996). The Cognitively Guided Instruction (CGI) team (Fennema et al., 1996; Franke, Carpenter, Fennema, Ansell, & Behrend, 1998; Franke et al., 2001) showed effects on (1) teacher knowledge and beliefs, (2) classroom instructional practices, and (3) student learning as a consequence of professional development that assisted elementary school teachers of mathematics in using more ambitious, student-centered teaching practices. Thus, studies at each grade band of school mathematics showed that ambitious instructional strategies improve student learning in mathematics. Why is it, then, that mathematics teachers do not consistently implement such teaching practices in their classrooms?

1.2.1 Possible barriers to ambitious mathematics teaching

While there are multiple possible causes for the lack of ambitious mathematics teaching, including multiple contextual factors like unsupportive administrators or communities that push for preservation of the status quo, no primary reason has been determined. Some possible explanations for why teachers do not employ ambitious teaching practices in mathematics may involve a lack of content knowledge for teaching (Ball, Thames, & Phelps, 2008; Shulman, 1986); a dearth of pedagogical content knowledge, especially in relation to effectively implementing the cognitively demanding mathematics tasks that are an essential part of ambitious instruction (Smith & Stein, 2011; Stein, Smith, Henningsen, & Silver, 2000); a set of beliefs and attitudes that are incongruent with teaching in an ambitious manner (Knapp & Peterson, 1995; Warfield, Wood, & Lehman, 2005); or preparation and on-going training that is not well-aligned with practices undertaken in the field (Grossman et al., 2009). This section will briefly discuss these possible reasons for the lack of ambitious mathematics teaching.

The first possible reason for lack of ambitious mathematics instruction mentioned above is a lack of content knowledge for teaching. According to Ball, Thames, and Phelps (2008), mathematical knowledge for teaching includes "knowledge of content and students…knowledge of content and teaching and…specialized content knowledge which is distinct from the common content knowledge needed by teachers and non-teachers alike" (p. 389). One of the most pervasive problems in mathematics education is the lack of content knowledge for teaching (Ball et al., 2008). So, is it plausible that a lack of mathematical knowledge for teaching influences the delivery of school mathematics to students such that ambitious mathematics teaching practices are either ignored or implemented without fidelity? What if teachers gained the necessary mathematical knowledge for teaching? Might their teaching practices become more ambitious in nature?

A second possible reason for the lack of ambitious mathematics instruction is a deficiency in pedagogical content knowledge (Ball et al., 2008; Shulman, 1986) related to effectively implementing cognitively demanding tasks. Examining one of the aspects of ambitious mathematics teaching as presented by Hiebert et al. (1997), that of using problematic, cognitively demanding tasks, makes it evident that teachers need either pre-service or in-service professional development to learn what such tasks look like and how to implement these tasks. Both the TIMSS video study (National Center for Education Statistics [NCES], 2003), and research findings from the QUASAR project (Stein et al., 1996), demonstrate that if American mathematics teachers choose to implement cognitively demanding mathematics tasks, they tend to lower the cognitive demand of such tasks when they are enacted. (NCES, 2003; Silver & Stein, 1996; Stigler & Hiebert, 1999). So, is it possible that attaining the pedagogical content knowledge to successfully implement cognitively demanding tasks can lay the groundwork for more ambitious mathematics instruction?

Ambitious mathematics teaching practices are meant to attend to and be responsive to student thinking (Huinker & Bill, 2017). Beliefs and attitudes incongruent with ambitious mathematics teaching practices provide a third possible reason why teachers do not employ ambitious teaching methods. Some mathematics teachers subscribe to the belief that there is such a thing as a "math brain," and that students either have it, or they don't (Boaler, 2015; Dweck, 2006). If teachers subscribe to belief in a "math brain," then student thinking about mathematics either gives the right answer or not, without any room for variation. Teachers, themselves, were often taught via rigid methods that emphasized one "right" or expected way of solving mathematics problems based on memorized facts and algorithms. Since teachers tend to teach in the way they were taught, learning to teach in their K-12 years as opposed to their teacher training courses (M. S. Smith, 2001; Wiliam, 2013), they may not have beliefs or attitudes conveying that all students are capable of doing mathematics to high levels (Boaler, 2013) or that there are many and varied ways to solve mathematics problems. If teacher attitudes and beliefs change to reflect that all students are capable of learning mathematics, might mathematics teaching become more ambitious in nature? If teachers begin to embrace the multiple ways mathematics problems may be solved, and become responsive to the student thinking behind the various methods, might there be evidence of that belief in their pedagogy[2]?

One final reason why mathematics teachers may not employ ambitious teaching practices comes from the work of Grossman and her colleagues (2009). They discussed three common

---

[2]. While teachers' beliefs and attitudes provide a viable reason for why ambitious mathematics teaching is lacking, the author will not explore this reason during the anticipated study.

approaches for training future clergy, teachers, and psychologists. All these developmental programs involved representation, decomposition, and approximation of practice as key pedagogical components. In teaching, examples of representations of practice include vignettes of classroom episodes or videos explaining particular teaching techniques, like wait time. Decompositions of practice in the teaching profession include writing learning goals for a lesson or rehearsing the process of giving directions. "Approximations of practice refer to opportunities to engage in practices that are more or less proximal to the practices of a profession" (p. 2056). In teacher training, approximations might be launching an actual classroom lesson with a mentor present or scripting a portion of a lesson as it is being taught. Unfortunately, teaching has the least well-developed approximations of practice of the pre-professional programs studied by Grossman and her colleagues. The majority of what is commonly part of teacher professional development is made up of representation and decomposition of practice. Close approximation of practice is not the norm in most professional development sessions (Grossman et al., 2009), as it becomes difficult to replicate realistic classroom experiences with only teachers in attendance. Does the lack of close approximation of practice during professional development account for why many teachers of mathematics do not employ ambitious teaching practices? If more close approximation of practice were included in continuing teacher training, might teachers' practice become more ambitious?

Much of this section alludes to a deficiency or need for change in mathematics teachers' knowledge, background, or experiences. Thus, it seems reasonable to conclude that teachers need professional development to deepen and expand their mathematical knowledge for teaching and pedagogical content knowledge (Ball et al., 2008) and to increase their opportunities to engage in decompositions, representations and approximations of practice (Grossman et al., 2009). The next section examines responses to the barriers to ambitious instruction described above.

1.2.2 Responses to barriers to ambitious instruction

Despite the shortfalls and possible reasons for a lack of ambitious mathematics teaching cited in the previous section, there *have* been studies of professional development that has been effective in changing teacher's classroom practice. While a lack of content knowledge for teaching provides one reason teachers do not employ ambitious practices, Heather Hill and colleagues have shown that there is a connection between teachers' professional development experiences and (a) their mathematical content knowledge for teaching (Hill & Ball, 2004), (b) their classroom pedagogy related to mathematics instruction (Hill, Blunk, et al., 2008) and (c) their students' subsequent mathematics learning (Hill, Rowan, & Ball, 2005). With these studies in mind, perhaps there is reason to believe that an increase in content knowledge for teaching can positively effect classroom practice to make it more ambitious.

Hill and colleagues investigated changes in classroom pedagogy as a result of professional development, but other studies showed an impact on specific aspects of mathematics classroom pedagogy as a result of professional development. The second reason given in this chapter for the lack of ambitious mathematics instruction is a deficiency in pedagogical content knowledge related to effectively implementing cognitively demanding tasks. The writings of Boston and Smith (2009, 2011) demonstrate that teachers *can* learn to implement high-level tasks while maintaining the level of cognitive demand inherent in the written version of the task. Boston and Smith worked with school-based mentor teachers for Masters of Arts in Teaching (MAT) students over two successive school years while examining the teachers' use of tasks (Boston, 2013; Boston & Smith, 2009). The researchers subsequently followed up with seven of the teachers from the original study, demonstrating retention of the use of high-level, cognitively demanding tasks after the professional development concluded (Boston & Smith, 2011). With these studies in mind, perhaps there is reason to think that, like increasing mathematical knowledge for teaching, learning to effectively implement cognitively demanding tasks translates to making classroom practice more ambitious.

A third possible reason provided for why teachers do not employ ambitious practices is that they may have beliefs and attitudes that are incongruent with teaching in an ambitious manner. Warfield, Wood, and Leham (2005) found that when teachers believed in their own and their students' autonomy or ability to make decisions for themselves, there was a greater tendency to encourage and support student-created solution strategies for novel problems. This potentially relates to several of the effective mathematics teaching practices like *eliciting and using student thinking* and *supporting productive struggle*, among others (NCTM, 2014). Additionally, researchers in the CGI studies found that when teachers believed that children's thinking was at the heart of their teaching practice, they were more likely to employ ambitious mathematics teaching (Franke et al., 1998, 2001). [3]

The last reason posed for why teachers do not employ ambitious mathematics teaching is preparation or on-going training that is not well-aligned with what actually occurs in the field. Cohen and Ball (1999) and Smith (2001) provide evidence that teacher training can align with what is undertaken by teachers in the field. They write that when teachers encounter materials from actual classrooms, akin to the close approximations of practice described by Grossman et al.

[3] While teachers' beliefs and attitudes is a viable reason for why ambitious mathematics teaching is lacking, the author will not explore this reason during the anticipated study.

(2009), they have the chance to examine, explore, critique, and simply think more about some new or different teaching practices. To illustrate, Smith cites the example of professional development in which teachers interact with a "mathematical task along with a carefully selected set of student responses" and argues for teachers to engage in "the work of teaching" (p. 8) by planning, enacting, and reflecting upon instruction during professional development activities. Cohen and Ball bring up the idea of using tasks that are specifically designed for teachers in order that they engage directly with the work done during teaching, even though they are in a professional development setting. These author-researchers seem to argue that when close approximations of the teaching practice are a part of professional development, then teachers have the opportunity to transfer that learning to their classroom and engage in more ambitious teaching.

While the counterarguments above show that there are cases of teachers using more ambitious teaching practices after: (1) increasing mathematical knowledge for teaching; (2) learning about the pedagogy of sustaining cognitive demand with task implementation; (3) adopting beliefs and attitudes that convey an interest in student learning; or (4) participating in professional development with close approximations of practice; these instances are certainly not ubiquitous. While the mathematics literature identifies cases where mathematics teachers learn to implement and sustain ambitious teaching practices (Boston & Smith, 2009, 2011; Franke et al., 1998, 2001), there does not seem to be a consistent connection between professional development and use of ambitious practices in mathematics education (Horizon Research Inc., 2013; Staples, 2007; Wang & Romero, 2013). Therefore, the need for teacher professional development or experiences that allow for close approximation of practice while increasing mathematical knowledge for teaching and increasing teachers' opportunities to learn about implementing cognitively challenging tasks with fidelity still exists. Perhaps stand-alone, outside-the-classroom professional development is not enough to consistently change teachers' methodologies in the classroom. Perhaps something more is needed to increase mathematics teachers' use of ambitious teaching practices.

1.3 The problem of practice

For more than 14 years, my work has focused on delivering professional development to PK-12 mathematics teachers, therefore, the connections between professional development and ambitious mathematics instruction are of interest. My workgroup, the Math & Science Collaborative (MSC), and I have experienced, firsthand, that even with professional development many teachers still do not implement ambitious teaching methods in their classroom. Over the years, my workgroup's mathematics institutes focused on the development of mathematical knowledge for teaching, as elucidated by Ball and her colleagues (2008; Hill et al., 2008); the development of ambitious mathematics teaching practices, including the use of cognitively-demanding, high-level tasks (Boston & Smith, 2011; Smith, Hughes, Engle, & Stein, 2009), the use of classroom discourse (Chapin, O'Connor, & Anderson, 2009; Michaels, O'Connor, & Resnick, 2008; M. S. Smith et al., 2009), and the use of appropriate mathematical tools (Carpenter, Fennema, Franke, Levi, & Empson, 1999; Hiebert et al., 1997; Huinker & Bill, 2017) among other ambitious practices. We also worked to help teachers develop positive beliefs and attitudes towards student-centered instruction (Warfield et al., 2005). Additionally, the professional development employed the key components of Grossman et al.'s (2009) pedagogies of practice, namely representation, decomposition, and approximations of practice.

In contrast to some of the previously cited work, where teachers employed ambitious teaching practices (Boaler & Staples, 2008; Boston & Smith, 2009; Franke et al., 2001), my group's work did *not* show similar results (Removcik, 2014; Romero & Winters, 2013; Wang, 2013; Wang & Romero, 2013). Few teachers used the pedagogy of ambitious teaching in their classroom or retained any ambitious teaching practices. Thus, the relevant questions become: Why is it that even with professional development geared towards ambitious mathematics instruction and with increased mathematical content knowledge, some teachers do not implement ambitious teaching methods in their classroom? Why do mathematics teachers continue with or return to traditional models of teaching? Do teachers need more or different experiences to help them bridge the gap between the professional development setting outside their classroom and the student-teacher interactions that take place in their classrooms? Could it be a lack of close approximations of practices, as Grossman and colleagues (2009) suggest, that account, at least in part, for the lack of ambitious teaching practices in use in mathematics classes? Might including professional development work inside the classroom help increase the use of ambitious mathematics teaching?

The proposed study aims to find out if more proximal support for teachers helps or catalyzes teachers' enactment of ambitious teaching practices that value and "attend to student thinking in an equitable and responsive manner" (Huinker & Bill, 2017, p. 4). More specifically, the purpose of this research study is to find out if pairing content-focused, outside-the-classroom professional development with coaching mathematics teachers in their classrooms impacts mathematics teaching to make it more ambitious. If so, how does that impact compare to the impact of the content-focused professional development without the added coaching component? Thus, the goal of the study is to investigate whether coaching added to outside-the-classroom professional development correlates with an increase in use of ambitious teaching practices.

1.4 Inquiry setting

Putnam and Borko (2000) offer that in order to situate learning experiences for in-service teachers within their practice, professional development might take place in their schools or even their classrooms. One possible method for increasing proximal support for teachers that involves greater use of in situ professional learning (Putnam & Borko, 2000) is content-focused coaching (Gibbons & Cobb, 2016; West & Staub, 2003). Sustained coaching, over a period of years, has been shown to increase student achievement in schools (Campbell & Malkus, 2011). The implementation of coaching within an adaptable model, wherein the coach can alter the implementation within given parameters, has also been shown to be effective (Russell et al., 2019).

Evidence exists to show that professional development can effect teacher practice (Boston & Smith, 2009, 2011; Franke et al., 1998) and that teacher practice effects student learning (Boaler & Staples, 2008; Fennema et al., 1996; Silver & Stein, 1996). Evidence exists to show that coaching can change teacher practice (Matsumura, Garnier, & Spybrook, 2013) and increase student achievement (Campbell, 2012; Campbell & Malkus, 2011). Thus, it seems the combination of coaching and outside-the-classroom professional development, both addressing ambitious mathematics teaching could be more effective than either outside-the-classroom professional development or coaching alone in changing teacher practice to make it more ambitious. Others have published about the combination of coaching and professional development. Neufeld and Roper (2003) wrote, “in light of our current knowledge about what it takes to change a complex practice like teaching, there are reasons to think that coaching, in combination with other professional development strategies, is a plausible way to increase schools’ instructional capacity” (p. 1). Cobb and Jackson (2011) concurred in calling for this pairing, saying what is needed is a “coherent system of supports for ambitious instruction” (p. 9) involving coaching as a key component for “improving mathematics instruction at scale” (p. 9). I submit that if the professional developer helps the in-service teacher decompose and represent his or her practice during the outside-the-classroom professional development experience *and* is present to help the teacher approximate their professional practice (Grossman et al., 2009) while in the classroom, the teacher will be more likely to employ ambitious mathematics teaching practices.

1.5 Inquiry questions

The compilation of my previous experiences and much of the research referenced within caused me to consider investigating how the proximity of a more knowledgeable other, acting as both the professional developer and a coach, in the classroom environment might change the level of implementation of ambitious mathematics teaching practices. I hope to discover whether more proximal support for teachers helps or catalyzes mathematics teachers’ enactment of ambitious teaching practices. The proximal professional development will include working with the teacher in their school and classroom to plan, teach, and reflect upon the lesson. This type of professional development is, in essence, content-focused coaching (Gibbons & Cobb, 2016; Matsumura, Garnier, Correnti, Junker, & Bickel, 2010; West & Staub, 2003).

I hypothesize that increasing the proximity of the professional development to the classroom will increase the likelihood of ambitious mathematics teaching practices. Therefore, the investigable research questions are:

How does proximal, in situ professional development in the form of content-focused coaching paired with outside-the-classroom professional development facilitate a change in mathematics teachers' pedagogical practices from traditional to more ambitious?

- *What is the impact on teachers'* opportunities to learn *about ambitious teaching practices when content-focused coaching is added to professional development?*

- *What is the impact on teachers'* use *of ambitious teaching practices when content-focused coaching is added to professional development?*

2.0 Literature Review

Ambitious mathematics instruction (Lampert et al., 2011) positively impacts student learning of the subject (Boaler & Staples, 2008; Carpenter & Fennema, 1991; Hiebert & Grouws, 2007; Stein & Lane, 1996). However, teachers of mathematics at K-12 often do not employ ambitious mathematics teaching pedagogies (Hiebert et al., 2005; Horizon Research Inc., 2013). Despite availability of professional development (PD) experiences aligned with adult learning needs (M. S. Smith, 2001), and research on effective professional practice (Darling-Hammond & Bransford, 2005; Garet, Porter, Desimone, Birman, & Suk Yoon, 2001), teachers often persist in employing more traditional methodologies in their classrooms or do not sustain changes made to their teaching methods (Removcik, 2014; Staples, 2007; Stein & Wang, 1988; Wang & Romero, 2013). The intention of this study is to uncover evidence about whether pairing content-focused coaching (Gibbons & Cobb, 2016; Matsumura et al., 2010; West & Staub, 2003) with outside-the-classroom PD catalyzes a change in teaching practice from traditional to ambitious. Before embarking on the study, the literature surrounding ambitious mathematics teaching as well as content-focused coaching needs investigating.

This chapter provides a literature review of the concepts relevant to the study. First, the chapter recounts a portion of the rich history of literature surrounding *ambitious instruction* and transitions to discussing current conceptions of ambitious mathematics teaching. Then, the chapter examines the literature showing that teachers of mathematics can change their practice from traditional to ambitious with PD supporting this change in practice. The last part of this chapter reviews the body of research on coaching in the classroom. The review includes an examination of coaching's roots, its evolution, and its current state within the research findings. The chapter closes by sharing some information relevant to the relationship between the proposed study and the current state of the literature on ambitious mathematics teaching and content-focused coaching, making the case that despite the extensive research base for ambitious instruction and despite the developing and evolving work being done in the field of coaching, there is room in the research field for this study.

2.1 The roots of ambitious teaching

There is a rich history surrounding what is currently called ambitious instruction in mathematics. What this dissertation calls ambitious mathematics instruction is defined by Lampert, Boerst, and Graziani as "teaching that aims to teach all kinds of students to not only know academic subjects, but also to be able to use what they know in working on authentic problems in academic domains" (2011, p. 1361). Mathematics instruction fitting this description has been labeled in a variety of different ways since its inception. Over the years, different projects and different researchers have used different labels for what this study calls ambitious mathematics teaching. The terms and labels used have evolved and, in some cases, become more precise.

2.1.1 Terminology associated with ambitious mathematics instruction

Some early terminology used to describe ambitious mathematics teaching fell out of favor with time and the introduction of different terminology. Members of the Quantitative Understanding: Amplifying Student Achievement and Reasoning (QUASAR) research team used the term *enhanced* mathematics instruction in some early research publications (Silver, Smith, &

Nelson, 1995; Silver & Stein, 1996; Stein et al., 1996) when referring to instruction involving the implementation of high-level mathematics tasks and the use of classroom discourse in association with task implementation. Silver and Stein (1996) also used the term *meaning-oriented* instruction (p. 503) in contrast to instruction emphasizing memorized procedures. *Meaning-oriented* instruction uses cognitively challenging tasks to encourage deeper understanding of the concepts which students are learning.

Other terminology used to describe what is now known as ambitious instruction employs the word "inquiry." *Inquiry-oriented* was used by Silver (1994) and Hughes, Smith, Boston, and Hogel (2008). Silver used the term in the context of mathematical problem posing by discussing how problem posing is a natural part of *inquiry-oriented* instruction, which also included student problem solving and "discovery" of mathematical ideas. While the term "discovery learning" is no longer prevalent, ambitious instruction does include student formation of conceptual understanding enabled by engagement with cognitively challenging problems. Hughes et al. used the term *inquiry-oriented* in connection with teacher professional development provided during the Enhancing Secondary Mathematics Teacher Preparation (ESP) program. In the ESP program, *inquiry-oriented* instruction entailed the use of cognitively-challenging mathematics tasks and the related teaching practices meant to sustain a high-level implementation of the task. These practices included supporting and using student thinking about the task and related mathematics via questions and discourse practices that specifically call for justifying strategies used. The term *inquiry-based* was used by Fraivillig, Murphy, and Fuson (1999). This team wrote about a framework devised to support teachers during inquiry-based mathematics instruction. The framework had three parts: "Eliciting Children's Solution Methods, Supporting Children's Conceptual Understanding, and Extending Children's Mathematical Thinking" (p. 148). The labels for the parts of Fraivillig et al.'s framework are reminiscent of the eight effective Mathematics Teaching Practices (NCTM, 2014) describing ambitious mathematics instruction.

The term "reform" was also frequently used to describe teaching practices that align with ambitious mathematics instruction. *Reform-oriented* was in the writings of Hiebert et al. (1997), Borasi and Fonzi (2002), and Boaler and Staples (2008). As mentioned in chapter 1 of this document, Hiebert and colleagues provided details around the elements of reform-oriented mathematics instruction in their publication, *Making Sense: Teacher and Learning Mathematics with Understanding* (1997). Components of reform-oriented instruction included changes in classroom tasks, the teacher's role, the nature of and way in which tools are used, the classroom culture, and the accessibility and equity in the classroom. Like Hiebert et al. (1997), Borasi and Fonzi (2002) provided a definition of "reform-oriented" mathematics teaching, when they wrote that it "involves much more than 'superficial features' such as using manipulatives …. Rather, …we refer to a comprehensive approach to mathematics instruction that is centered on teaching for understanding and enabling students to engage with meaningful problems and 'big ideas'" (p. 9). Paramount in this definition were (1) developing student understanding, and (2) confronting meaningful mathematics by way of problems encountered. Boaler and Staples (2008) also referred to *reform-oriented* mathematics instruction. While the authors did not provide a definition for *reform-oriented* instruction like Hiebert et al. (1997) or Borasi and Fonzi (2002), they did provide that the "demands placed upon students in reform-oriented classrooms are quite different than those in more traditionally organized classrooms" (p. 610) and success of such an approach "depends on teachers' careful and explicit attention to the ways students may be helped to participate in new learning practices" (p. 611).

Other terms employing the word “reform” to describe ambitious mathematics teaching are *reform-aligned* and *reform practices* both used by Staples (2007), and *reformed teaching* used by Sawada et al. (2002). Staples discussed practices such as questioning, communicating, explaining, and sense-making. Sawada and colleagues provided at least a partial definition for *reformed teaching* as “a movement away from the traditional didactic practice…Reform presupposes that teachers do not emphasize lecture, but rather stress a problem-solving approach and foster active learning” (p. 246).

Finally, a term used quite often to imply ambitious teaching is *Standards-based instruction*. A number of researchers employed this term over the last few decades, including but not limited to, Tarr, et al. (2008) and Stein, Smith, Henningsen, and Silver (2000; 2009). Tarr and colleagues used the term in connection with a type of learning environment aligned to the view of student learning apparent in the NSF-funded curricular materials. Tarr et al. labeled this classroom environment a Standards-Based Learning Environment (SBLE) and described the SBLE as one that generally encourages “active engagement of students, a focus on problem-solving, and attention to connections between mathematical strands as well as real-life contexts” (p. 248). Stein et al.’s (2000, 2009) publication, *Implementing Standard-Based Mathematics Instruction,* was about using–choosing, setting-up, and carrying out–high-level, cognitively challenging mathematical tasks in the classroom. Stein et al.’s text discussed mathematics tasks, their cognitive demand, which is explained as “the kind and level of thinking required of students in order to successfully engage with and solve the task” (p. 1), and their classroom implementation. The text then employed a series of cases, to illustrate “research-based patterns of teaching and learning” (p. xxi) and help readers connect their own classroom instructional patterns, to the cases in hopes of supporting the use of *Standard-based* instructional strategies.

Use of the term *Standards-based instruction* leads to the NCTM Standards documents (1989, 1991, 1995, 2000b), especially the *Professional Standards for Teaching Mathematics* (NCTM, 1991). This document contained six standards for mathematics instruction from which the term *Standards-based* is derived. These standards were presented in four sections on *tasks, discourse, environment,* and *analysis.* These standards had labels such as: *Worthwhile Mathematical Tasks*; *Teacher's Role in Discourse*, *Learning Environment* and *Analysis of Teaching and Learning.* The sections explaining each of the standards within the *Professional Standards* document contain verbiage much like that in other NCTM documents that were yet to be published in 1991, such as the *Principles and Standards for School Mathematics* (NCTM, 2000) and *Principles to Action: Ensuring Mathematical Success for All* (NCTM, 2014).

Borko, Kieran, and Lester (2004), as quoted in Tarr et al. (2008), observed that "the general mathematics education community too often uses terms such as 'standards-based instruction,' 'reform-based classrooms,' 'problem-based instruction,' and 'inquiry-based teaching' interchangeably" (p. 266). This statement seems to imply that the terms are not similar enough to be interchangeable. However, thorough examination of the *Professional Standards* (NCTM, 1991) leads to the conclusion that the other terminology mentioned in this section and associated with ambitious mathematics instruction is all connected to a core set of teaching practices that are similar in multiple facets, such as: emphasis on problems, high-level tasks, or challenging classroom experiences in mathematics; use of classroom discourse; and envisioning a classroom where all students' contributions are valued.

2.1.2 Early calls for ambitious teaching

While some of the roots of ambitious mathematics teaching are in the *Standards* documents (NCTM, 1989, 1991, 1995, 2000b), another prevalent starting point for research around ambitious mathematics teaching lies with Lee Shulman's (1986) Presidential address at the American Educational Research Association (AERA) annual meeting. Shulman began a conversation that continues today by stating that the practice of teaching needed to change and improve. He put forth a new theoretical framework around the practice of teaching. Shulman said more study about knowledge types and skills needed for teaching was necessary, and he proposed three categories of teacher knowledge: subject matter content knowledge, which is "the amount and organization of knowledge per se, in the mind of teachers" (p. 9); curricular knowledge, which includes understanding how and why topics are arranged a certain way in the curriculum; and pedagogical content knowledge, which is "subject matter knowledge for teaching" (p. 9).

This was the first mention of a category of teacher knowledge blending content and pedagogy. Shulman bolstered his argument for its existence by stating,

> Mere content knowledge is likely to be as useless pedagogically as content-free skill. But to blend properly the two aspects of a teacher's capacities requires that we pay as much attention to the content aspects of teaching as we have recently devoted to the elements of teaching process. (1986, p. 8)

In addition to introducing educational researchers to the idea of pedagogical content knowledge (1986, 1987), Shulman also inspired a line of research into training teachers in the practice of teaching when he specifically called for case study development and usage in training programs; a call that continued in other writing (Shulman, 1998). Through these writings and by attempting to frame and codify effective teaching in content areas, Shulman and his team established a foundation for the continuing conversation about ambitious teaching practices and related constructs.

2.1.3 Other constructs are related to ambitious instruction

2.1.3.1 Ambitious instruction leads to increased student learning

Shulman's (1986) address, along with NCTM's *Standards* documents (1989, 1991, 1995, 2000b) and government funding from the National Science Foundation (NSF), prompted multiple lines of research in mathematics education. Some research examined classroom instruction and its impact on student learning, showing that ambitious mathematics instruction increased student achievement (Boaler & Staples, 2008; Carpenter, Fennema, Peterson, Chiang, & Loef, 1989; Stein & Lane, 1996). As discussed in chapter 1 of this dissertation, ambitious teaching practices are connected to increased student learning in the mathematics classroom at every grade band.

Cohen and Ball (2001) and Stein, Remillard, and Smith (2007) showed that it is the instruction, not textbooks or curricular materials, that makes the difference in student learning. Multiple studies reinforce this finding. Research associated with Cognitively Guided Instruction (CGI) showed that ambitious mathematics instruction at the elementary level leads to increased student achievement (Carpenter et al., 1989; Fennema et al., 1996). The QUASAR project's interrelated studies (Henningsen & Stein, 1997; Silver & Stein, 1996; Stein et al., 1996; Stein & Lane, 1996; Stein & Smith, 1998) done at the middle school level indicated that "mathematical tasks with which students become engaged determine not only what substance they learn but also how they come to think about, develop, use, and make sense of mathematics" (Stein et al., 1996, p. 459). Furthermore, "the greatest learning gains for students are realized when students have consistent opportunities to engage with high-level tasks" (M. S. Smith & Stein, 2018, p. 130).

Boaler and Staples' (Boaler, 2006a, 2006b; Boaler & Staples, 2008) longitudinal study of the pedagogical practices in high school mathematics classes evidenced that ambitious instruction increased student learning in high school, showing that students in a de-tracked, urban-like high school setting significantly outperformed their peers in tracked and suburban-like settings when ambitious teaching practices were employed. While there are differences among the studies, including but not limited to the grade bands, the CGI studies, the QUASAR studies, and the Railside study all demonstrate one common finding: Ambitious mathematics practices positively impact student achievement.

2.1.3.2 Increased teacher knowledge leads to ambitious instruction

Another line of research directly influenced by Shulman's (1986) challenge addressed the types of teacher knowledge, skills, proficiencies, expertise, and capacities needed to effectively instruct in mathematics classrooms. While there are those in the mathematics community, like Askey (2001), who contend that mathematics teachers in the United States do not have the content knowledge needed to teach in an ambitious manner, others, like Ball and her colleagues (Ball & Bass, 2000; Ball et al., 2008), believe there is more to teaching than extensive content knowledge. They examined subdomains of content knowledge and other forms of teacher knowledge, which have their genesis in the writings of Shulman (1987). The work of Ball and Bass (2000) on bridging content and pedagogy led the way to Ball, Thames, and Phelps' (2008) conceptualization of the ideas into the domains of *Mathematical Knowledge for Teaching* (MKT) which is "the mathematical knowledge needed to carry out the work of teaching mathematics" (p. 395). (See Figure 2.1.) Mapping the domains of MKT was a large step towards understanding what teacher knowledge makes a difference for ambitious classroom practice.

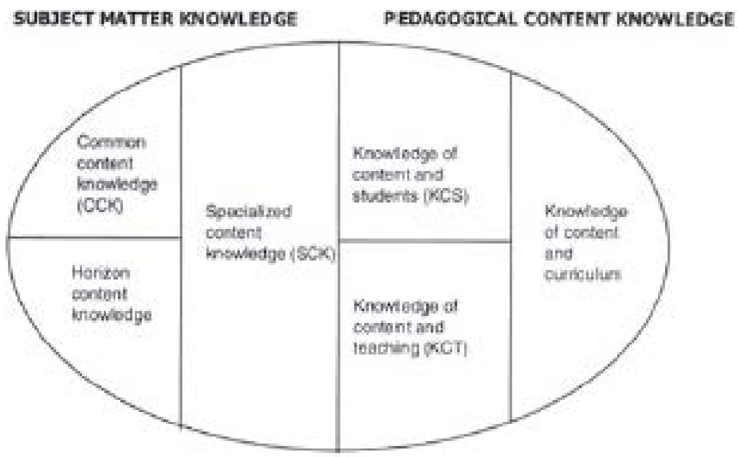


Figure 2.1 Domains of Mathematical Knowledge for Teaching (Ball et al., 2008, p. 403)

An additional consequence of the research done by Ball's Learning Mathematics for Teaching (LMT) team was the eventual development of a set of survey instruments that reliably and validly measured the component parts of mathematics teachers' knowledge for teaching[4] (Hill & Ball, 2004; Hill, Ball, & Schilling, 2008; Hill, Shilling, & Ball, 2004). Twenty years after his initial address, this team brought additional clarity, detail, and coherence to Shulman's (1986, 1987) ideas and conceptual framework regarding teacher knowledge.

Boston (2013) undertook a study related to increased teacher knowledge within the Enhancing Secondary Mathematics Teacher Preparation (ESP) project. Boston's study explored whether a change in teacher knowledge took place as a result of ESP PD. Boston used Desimone's (2009) conceptual framework as a model for tracing the effects of the PD through changes in teacher knowledge regarding high-level tasks to changes in teacher practice with regard to

[4] The paper-and-pencil assessment of MKT, called the Learning Mathematics for Teaching (LMT) assessment, was used as a means of matching coached and comparison groups before the start of the study for this dissertation. That was the extent of its use. It was not used to re-assess teacher MKT after coaching, and teacher knowledge was not used as a variable in this study.

selecting and implementing high-level tasks. The study gathered data on teacher knowledge via a written-response task-sort, and the data was connected to learnings from the PD. The participants in the PD significantly increased their scores from pre- to post-PD, showing increased teacher knowledge regarding cognitive demands. At the conclusion of the ESP PD, participants also had significantly more knowledge than the contrast group of secondary teachers who has not participated in the ESP PD. Although Boston (2013) did not employ the tools developed by Ball's group to measure teacher knowledge, she surmised one possible reason for the change in teachers' ability to successfully select and implement cognitively challenging tasks was teachers' increased knowledge.

Aside from research already reviewed, the initial roots of ambitious instruction provided by Shulman (1986) and the *Standards* (NCTM, 1989, 1991, 1995, 2000b) inspired other lines of research related to ambitious mathematics instruction. One research line examined the impact of PD on content knowledge for teaching (Bell, Wilson, Higgins, & Mccoach, 2010; Hill & Ball, 2004). Another, which will be reviewed later in this document, examined the impact of PD on mathematics instruction, showing that teachers' practice *can* be positively impacted by professional development (Boston & Smith, 2009, 2011; Franke et al., 1998; Knapp & Peterson, 1995). Other teams influenced by Shulman examined the possible connection between teacher knowledge and student learning (Ball, Hill, & Bass, 2005; Hill et al., 2005). Still others worked on tools to qualitatively and quantitatively measure ambitious mathematics instruction (Learning Mathematics for Teaching Project, 2011; Matsumura et al., 2013; Sawada et al., 2002). These tools will also be reviewed next, as they comprise one last construct related to ambitious instruction.

2.1.3.3 Ambitious mathematics teaching can be measured

As previously stated, for decades, studies at elementary, middle, and high school have provided evidence that ambitious mathematics instruction results in increased student learning (Boaler & Staples, 2008; Carpenter et al., 1989; Stein & Lane, 1996). Some of the studies demonstrating increased student learning employed classroom observation tools to help make the link between ambitious instruction and student learning. For example, Fennema et al. (1996) used a tool measuring “Levels of Cognitively Guided Instruction” (p. 412) in their study with first- through third-grade teachers. The QUASAR project created their own Classroom Observation Instrument (COI) and used it to evaluate the implementation of tasks (Henningsen & Stein, 1997; Stein et al., 1996). Currently, there are three tools that research studies frequently use for examining instruction in mathematics classrooms: the Mathematics Quality of Instruction (MQI) tool; the Reformed Teaching Observation Protocol (RTOP); and the Instructional Quality Assessment (IQA).

The Mathematical Quality of Instruction (MQI) classroom observation tool links teacher knowledge to classroom practice. While the Learning Mathematics for Teaching (LMT) team developed paper-and-pencil measures of a subset of domains related to Mathematical Knowledge for Teaching (MKT) (Hill & Ball, 2004; Hill, Ball, et al., 2008; Hill et al., 2005), they also worked to develop an observational instrument for use in classrooms. Hill, Blunk, et al. (2008) published an exploratory study which showed a relationship between teachers’ MKT, measured via the paper-and-pencil survey, and their classroom instruction, assessed via the MQI. Teachers who scored higher on the paper-and-pencil assessment measuring MKT exhibited a number of ambitious teaching practices including an insistence on mathematical explanations of student thinking, use of discourse moves like agreeing or disagreeing with classmates’ reasoning, using multiple representations or multiple solution methods, and choosing and sequencing mathematical tasks for instruction. Generally, teachers with low-MKT, exhibited fewer ambitious mathematics teaching practices than their colleagues with high MKT. "In terms of both affordances and deficits, high-MKT teachers provide better instruction for their students" (Hill, Blunk, et al., 2008, p. 457).

The MQI has seven scales (Hill, Blunk, et al., 2008) with multiple subscales for each (LMT, 2011). During the viewing of a videotaped lesson, scoring for the subscales occurs at five-minute intervals as "present-appropriate," "present-inappropriate," "not present-appropriate," or "not present-inappropriate." The score for each of the seven scales is obtained by averaging the number of "present-appropriate" or "not present-appropriate" scores, so the MQI can be cumbersome to score. The MQI instrument was revised in 2014 (Boston, Bostic, Lesseig, & Sherman, 2015). Codes were refined and the instrument was made to "explicitly align with the mathematical practices outlined in the Common Core State Standards for Mathematics" (Boston et al., 2015, p. 161). This revised instrument emphasizes the use of in-context tasks and classroom discussion. Some subscales measure "student engagement in sense-making as indicated by the quality of student explanations; evidence of students' questioning, conjecturing, and generalizing mathematical ideas; and the cognitive demand of the task" (p. 161). Thus, descriptions of items in the subscales align with ambitious instruction.

The Reformed Teaching Observation Protocol (RTOP) (Sawada et al., 2002) is another instrument used to rate classroom mathematics (and science) lessons. It is a 25-item observation protocol for use in K-20 classrooms. The RTOP has three sections: Lesson Design and Implementation; Content; and Classroom Culture. The Content and Classroom Culture sections have two subscales each. With five items in each of the five sections rated on a 5-point (0-4) Likert-scale, the researchers were aiming for high internal consistency and ease of use. The highest possible overall score is 100, with an overall score of 50 or greater considered the minimum for a lesson having elements of reformed teaching. A score of 10 or greater on any subscale means there is evidence of reform orientation. Examination of sub-scores reveals if the reform orientation is consistent across subscales (Boston et al., 2015). Sawada et al. (2002) showed there was a correlation between RTOP scores and student achievement as measured by a comparison of pre- and post-tests on class content. As stated by the authors, “Data show that when teaching is highly reformed, student learning is significantly enhanced” (p. 251).

A final tool being reviewed in this chapter is the Instructional Quality Assessment (IQA) toolkit. “The IQA assesses elements of ambitious instruction in mathematics, specifically, the level of instructional tasks and task implementation, opportunities for mathematical discourse, and teachers’ expectations” (Boston, 2012a, p. 76). The IQA is meant to be used at scale without any required videotaping. It is based on the constructs of *Academic Rigor* (AR) and *Accountable Talk* (AT) and uses a combination of observations, assignments collections, and student work (Boston et al., 2015). The theoretical framework for the AR rubrics is the Mathematics Task Framework (Stein et al., 2009). The foundation for the AT rubrics comes from Resnick and Hall (1998) and is based on accountability to the community and to the discipline. The IQA mathematics toolkit uses a series of descriptive rubrics to assign a score (0-4) to elements of classroom instruction (Matsumura, Garnier, Slater, & Boston, 2008). Eleven of the rubrics in the IQA mathematics toolkit are connected to classroom observations. (See Appendix A.1 for a summary of rubric categories and titles used for classroom observations in various publications.) The remaining six rubrics in the IQA mathematics toolkit are connected to assignments given by the teacher and the student work done for those assignments. Using the collection of assignments and student work is one thing that makes this instrument unique. (See Appendix A.2 for a summary of rubric categories and titles used for assignment collections in various publications.) A score of 0 on a rubric indicates the item being measured is not present. Scores of 1 or 2 are considered low on the IQA. Low scores indicate the element is included but is of low quality and/or in low quantity with respect the desired state. To earn a high score of 3 or 4 on the IQA, the desired element must be present and be of high quality (Boston, 2012b, 2012d). A limited number of the IQA rubrics are purely holistic, but they all have a descriptive element. So, although the IQA scoring is quantitative and provides statistical data, there is a qualitative component to the tool (Boston, 2012a; Boston & Wilhelm, 2017).

Resnick, Matsumura, and Junker (2006), and Matsumura, Garnier, Slater, and Boston (2008) found that as few as four assignment collections or two classroom observations provided a "stable estimate of teacher quality" (Resnick et al., 2006, p. 1). Wilhelm and Kim (2015) performed a multivariate analysis and concluded that three or more observations are needed to reliably measure instructional quality with the IQA. Researchers found that both assignment collection and classroom observation were associated with elements of students achievement (measured via the SAT-10) (Boston, 2012a; Matsumura et al., 2008). Additionally, research established a strong association between the rubrics for classroom observations and those for assignment collections, demonstrating that, if observations are too time or resource intensive or too intrusive, collecting assignments and associated student work can stand in for classroom observations in mathematics classrooms. (Boston, 2012a; Matsumura et al., 2008; Resnick et al., 2006). The goal for the creators of the IQA toolkit was to robustly measure instructional quality without much burden for the teacher (Resnick et al., 2006). The two approaches used with the IQA–classroom observations and assignments collections with student work–directly measure enactment of content in the classroom, which is qualitatively different than using surveys or even teachers' instructional logs

(Matsumura et al., 2008). With all these qualities, the IQA is unique among the observational tools reviewed. However, the IQA does have some drawbacks. It does not assess teachers' mathematical accuracy in lesson delivery, like the MQI does, and because no videotaping is required, there is no permanent record of the instruction, aside from field notes. In summary, the IQA fills what was previously a gap in information about student learning by providing a "direct assessment of students' opportunities to learn mathematics" (Boston & Wilhelm, 2017, p. 833). Much of the aforementioned research, findings, and tools from the related studies have led to the current conceptions of ambitious instruction. This chapter now examines those current conceptions.

2.2 Current conceptions of ambitious instruction

Years prior to the publication of the CCSSM (NGA and CCSSO, 2010), *Adding it Up* (National Research Council, 2001) explained ambitious instruction as teaching aimed at ambitious learning goals for students. The publication said that in mathematics education, ambitious teaching has the goal of mathematical proficiency for all, where this proficiency involves the interconnected strands of conceptual understanding, procedural fluency, strategic competence, adaptive reasoning, and productive disposition. As time moved forward after the publication of CCSSM and the Next Generation Science Standards (NGSS) (NGSS Lead States, 2013), others contributed to defining high-quality, ambitious instruction. Lampert, Boerst, and Graziani (2011) wrote "'Ambitious teaching' is teaching that aims to teach all kinds of students not only to know academic subjects, but also to be able to use what they know in working on authentic problems in academic domains" (p. 1361). Cobb and Jackson (2011) wrote, "A central goal of ambitious teaching is that learning opportunities are distributed equitably (Lampert & Graziani, 2009;

NCTM, 2000). In this context, equity implies that all students should be able to participate substantially in all phases of classroom activities" (p. 8). More recently, Huinker and Bill (2017) wrote "Ambitious mathematics teaching involves skilled ways of eliciting and responding to each and every student in the class so that they learn worthwhile mathematics and come to view themselves as competent mathematicians (Anthony et al, 2015)" (p. 46). Each of these definitions of ambitious teaching involves all students and goes beyond viewing students' learning as reproducing facts or methods.

2.2.1 Ambitious instruction across content areas

Ball and Forzani (2009), in their article advocating for making teaching practice the core of teachers' professional preparation, outlined the component parts of ambitious teaching and labeled these as *high-leverage practices*. The work of Ball and Forzani's team at the University of Michigan in specifying high-leverage teaching practices across content areas resulted in, the creation of the TeachingWorks group and the related website (http://www.teachingworks.org/). This website defines a high-leverage practice as simply "an action or task central to teaching" ("The work of teaching," n.d., para. 2). While this website is not specifically aimed at mathematics teaching, it certainly includes practices applicable in the teaching of mathematics. For example, the first listed high-leverage practice on the website is "Leading a group discussion" ("High-Leverage Practices," n.d., para. 1), which is a process of talking, listening, and using the contributions of others to develop a better understanding of the content being addressed. The TeachingWorks group has thus far enumerated 19 high-leverage practices.

While work of the TeachingWorks group cuts across all content domains, another group, led by Windschitl and his colleagues at the University of Washington, seeks to apply ambitious teaching specifically to K-12 science. Their website, called "Tools for Ambitious Science Teaching" (https://ambitiousscienceteaching.org/), states "Ambitious teaching deliberately aims to support students of all backgrounds to deeply understand science ideas, participate in the activities of the discipline, and solve authentic problems" ("What is ambitious teaching," n.d., para. 1). The group created an Ambitious Science Teaching Framework comprised of four core sets of practices: "Planning for engagement with important science ideas; Eliciting students' ideas; Supporting on-going changes in student thinking; and Pressing for evidenced-based explanation" ("What is ambitious teaching," n.d.). Together these four sets of practices create a framework for science teaching that is different from traditional science teaching. Teaching science using Windschitl's framework is more ambitious.

2.2.2 Ambitious mathematics instruction

Staples (2007) wrote about the need for a better vision of what teachers should be doing in classrooms to enact a reform agenda. Ball and Forzani's group and Windschitl's group provided frameworks for teacher actions in line with an ambitious teaching agenda, thus providing a better vision. In the content area of mathematics, NCTM created an initial vision for ambitious instruction in mathematics via their standards documents (1989, 1991, 1995, 2000b). NCTM's vision was influenced by research cited throughout this chapter as well as research in cognitive psychology, social constructivism, and other academic arenas. Cobb and Jackson (2011) drew upon NCTM's vision to ground the learning goals of the districts in the *Middle School Mathematics and the Institutional Setting of Teaching* (MIST) project, and wrote:

> [NCTM's] vision is often referred to as ambitious teaching (Lampert, et al., 2010). In this vision, teachers support students to solve cognitively-demanding tasks (Stein, Smith,

Henningsen, & Silver, 2000), press students to provide evidence for their reasoning and to make connections between their own and their peers' solutions (McClain, 2002), and orchestrate whole class discussions in which they build on students' contributions to achieve their mathematical agendas for students' learning (Franke et al., 2007; Stein, Engle, Smith, & Hughes, 2008). Instructional practices of this type contrast sharply with typical teaching in most US classrooms and require teachers to anticipate and respond to students' thinking (Kazemi, Franke, & Lampert, 2009). (p. 8)

2.2.2.1 Principles to action: Ensuring mathematical success for all

NCTM's current vision of ambitious mathematics instruction was published in *Principles to Action: Ensuring Mathematical Success for All*, (NCTM, 2014) and took the form of the eight effective Mathematics Teaching Practices. "These eight Mathematics Teaching Practices… represent a core set of high-leverage practices and essential teaching skills necessary to promote deep understanding of mathematics" (NCTM, 2014, p. 9). As written in *Principles to Action*, "an excellent mathematics program requires effective teaching that engages students in meaningful learning through individual and collaborative experiences that promote their ability to make sense of mathematical ideas and reason mathematically" (NCTM, 2014, p. 7). These practices help the mathematics education community frame and concretize what an excellent mathematics program enacting ambitious teaching looks, sounds, and feels like. The effective Mathematics Teaching Practices are a succinct way to frame ambitious teaching in mathematics.

These teaching practices advocate for students engaging in challenging, collaborative work provided by the teacher and for teachers supporting learners in this work through their teaching actions, moves, and routines. The eight practices are: Establish mathematics goals to focus learning; Implement tasks that promote reasoning and problem solving; Use and connect mathematical representations; Facilitate meaningful mathematical discourse; Pose purposeful questions; Build procedural fluency from conceptual understanding; Support productive struggle in learning mathematics; and Elicit and use evidence of student thinking. While these practices "provide a framework for strengthening the teaching and learning of mathematics" (Huinker & Bill, 2017, p. 4), they do not provide the whole picture of ensuring mathematical success for all students. The *Principles to Action* (NCTM, 2014) book was constructed around five Guiding Principles for School Mathematics (p. 5) updated from the initial set of Principles for School Mathematics in *Principles and Standards for School Mathematics* (NCTM, 2000b). While the eight effective Mathematics Teaching Practices make up the guiding principle of Teaching and Learning, the Teaching and Learning principle is just one of the five Guiding Principles for School Mathematics. The other four guiding principles–Access and Equity, Curriculum, Tools and Technology, Assessment, and Professionalism–round out the *Principles to Action* book and give additional context to the eight effective teaching practices for mathematics.

2.2.2.2 Taking action: Implementing effective Mathematics Teaching Practices

Principles to Action (NCTM, 2014) provided initial information about each of the eight effective Mathematics Teaching Practices. For each practice, it offered some discussion of the practice including relevant research findings, an illustration of the practice with a classroom-based example, and a set of teacher and student actions indicative of the practice in use in the classroom setting. Three years hence, NCTM published a series of three texts, entitled, *Taking Action: Implementing Effective Mathematics Teaching Practices* (M. S. Smith, 2017), with one text aimed at each grade band: K-5, 6-8, and 9-12. These texts provided more information about how to successfully implement the eight effective teaching practices for mathematics

Each text contained ten chapters; one chapter for each of the eight practices. Spread throughout these chapters was a set of thinking exercises called Analyzing Teaching and Learning (ATL) activities. Each ATL prompted the reader to consider particular aspects of the effective teaching practice that was the focus of the chapter. The first chapter of each *Taking Action* book served to set the stage for ambitious instruction with a classroom vignette based on a grade band appropriate, cognitively demanding task. The same task or classroom episode was revisited in multiple chapters throughout the text to illustrate multiple effective teaching practices. The concluding chapter served to make the coherence and interconnectedness of the eight effective teaching practices for mathematics more explicit for the reader. Guiding the last chapter was a teaching framework showing the relationships in the practices. This framework is shown in Figure 2.2. The figure served as an illustration that while each of the practices contributes to ambitious mathematics teaching, ambitious mathematics teaching is more than simply thinking about each practice individually. In relationship to ambitious mathematics teaching, one must consider the whole set of practices as greater than the sum of the parts.

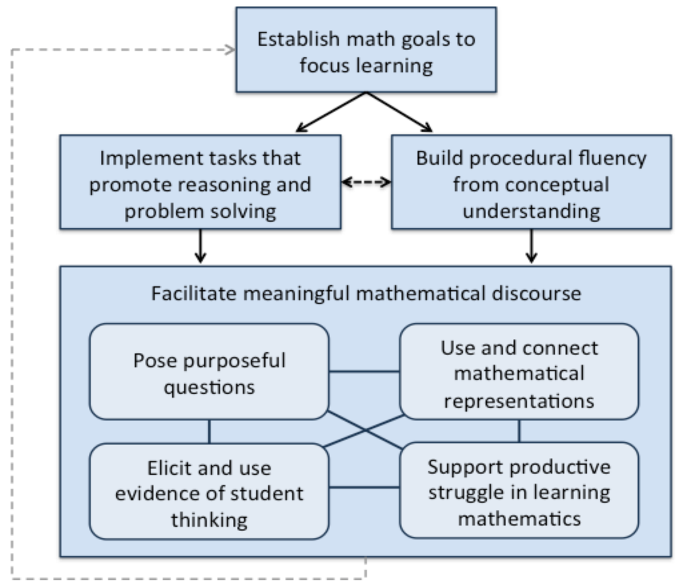


Figure 2.2 The eight effective Mathematics Teaching Practices shown in a framework highlighting the relationships between and among them (Huinker & Bill, 2017, p. 245)

NCTM continues to create tools for those embracing ambitious mathematics teaching, like the recently published update to the *5 Practices for Orchestrating Productive Mathematics Discussions* (M. S. Smith & Stein, 2018), but despite the tools and research, the multiple labels, and the refinement of the vision, the original NCTM vision explicated in the *Standards* documents (1989, 1991, 1995, 2000b) and furthered in the CCSSM (NGA and CCSSO, 2010), has not come to fruition in American classrooms (Horizon Research Inc., 2013; Stein et al., 2007; Stigler & Hiebert, 2004; Weiss, Pasley, Smith, Banilower, & Heck, 2003). Therefore, the need still exists for teacher PD that furthers the vision of those who have long advocated for reformed, standards-based, ambitious mathematics instruction. The chapter now turns to examining professional development that supports ambitious instruction in mathematics classrooms.

2.2.3 Outside-the-classroom Professional Development can support ambitious instruction

This chapter has discussed the evolution of terminology associated with ambitious mathematics instruction, early calls for ambitious instruction in the education community, and multiple studies showing that ambitious mathematics instruction results in greater student learning. The chapter has also discussed more recent conceptions of ambitious instruction as well as a few tools that will measure the quality of such instruction. This review now turns to literature showing the effects of professional development on teachers' classroom practice.

2.2.4 Task-centric professional development allows for changes in teacher practices

Boston and Smith (2009, 2011) and Boston (2013) undertook studies related to the ESP project. One component of the ESP project was a professional development initiative for secondary mathematics teachers who would later mentor a pre-professional mathematics teacher. Drawing on research from QUASAR which recognized the central role of high-level tasks in ambitious mathematics instruction (Henningsen & Stein, 1997; Stein & Lane, 1996), the PD was framed around the practice of using–choosing, setting up, and implementing–cognitively challenging classroom mathematics tasks. Following Shulman's (1986) recommendation, ESP PD used "case methods" (Stein et al., 2009, p. 23) to relate the implementation of the mathematics tasks under consideration to actual classrooms and to motivate teacher reflection on the tasks, the cases, and their own instruction.

Boston and Smith's (2009) study examined the effects of participation in the ESP PD on the selection and subsequent use of cognitively demanding tasks. The study considered task selection and implementation patterns for 18 participating teachers. Data consisted of five consecutive days' worth of instructional tasks and teacher log sheets, class sets of student work for three tasks used in this five-day period, and a classroom observation conducted during the same five-day period. Data were collected at three different junctures throughout the school year (Fall, Winter, Spring). The same data were collected for a contrast group of ten teachers in the Spring. Data were coded using the Instructional Quality Assessment (IQA) Academic Rigor (AR) rubrics for Potential of the Task and Implementation of the Task. The rubrics, which are on a 5-point scale (0 to 4; 0 meaning not present), consider ratings of 1or 2 as low-level and ratings of 3 or 4 as high-level. Rubrics were applied to the collected tasks, student work, and the classroom observation. Boston and Smith wanted to determine if teachers' changed their instructional practices around the (1) use and (2) implementation of tasks during and after the PD as compared to their own instruction before the PD and as compared to the contrast teachers. The study also wanted to (3) determine if the curriculum type (conventional or standards-based) influenced the use of cognitively challenging tasks.

The results showed that teachers who participated in the ESP PD significantly increased the average level of cognitive demand of the tasks selected for classroom use (i.e., mean score on the AR rubric for Potential of the Task) between the Fall and Winter and between the Fall and Spring. These gains were not influenced by the curriculum type being used. Participating teachers also significantly increased the *percentage* of high-level tasks selected, meaning more tasks with ratings of 3 or 4 were selected, when comparing Fall to Spring. Examining scores for *implementation* of the task yielded significant increases for the student work samples; however, scores for the classroom observations, while yielding higher scores, did not show significant increases on the Implementation of the Task rubric when participating teachers' scores were compared in Fall, Winter, and Spring. Participating teachers did score significantly higher than their counterparts in the contrast group for both *selection* and *implementation* of tasks. Boston and Smith discuss the implications of these results, stating, “These instructional changes…suggest that the ESP workshop can serve as one model of the type of professional development capable of supporting improvements in teachers’ instructional practices and students’ learning” (p. 147).

Boston and Smith (2011) performed a follow-up study with seven of the 18 teachers from the original ESP cohort to determine if the teachers sustained high levels for task selection and implementation. The researchers visited teachers’ classes more than a year after the conclusion of PD. Like the original study, the follow-up study used IQA AR rubrics to score tasks, student work, and lesson observations. Results showed that the subset of ESP project teachers participating in the follow-up study maintained the changes they had made in the original study by (1) continuing to *select* high-level cognitively challenging tasks for use in their classrooms, and (2) continuing high-level *implementation* of tasks. In fact, the percentage of teacher-chosen high-level tasks increased for this follow-up compared to the time period directly after PD.

Boston and Smith (2009, 2011) showed changes in the ability of participating teachers to select and implement cognitively challenging tasks and sustain their changed practices over time. Boston (2013) showed a connection between increased knowledge related to challenging tasks and classroom practice. She posed a hypothesis regarding this chain of events stating, “Teachers selected significantly more high-level tasks for instruction after their experiences in the workshop because they learned to attend to and value the opportunities for students’ learning embodied in such tasks” (p. 28). Results from the ESP project show that with under 40 hours of “task-centric” PD teachers can (1) change their knowledge about ambitious teaching practice, (2) change their practice from more traditional to more ambitious, and (3) sustain this change over time.

2.2.5 Focusing on children's thinking allows for changes in teacher practices

CGI literature showing increased student achievement as a result of teachers' implementing ambitious mathematics instruction has been reviewed. The PD component of CGI is considered now as an influence on changes in teacher practice. The PD associated with the CGI program was different from most mathematics workshops teachers attended at the time. The researchers did not train teachers in a new method of teaching. Rather, they shared (1) research findings showing that young children can solve many types of arithmetic word problems using a variety of materials and strategies; and (2) "frameworks" developed in conjunction with research (Carpenter et al., 1999). One framework categorized arithmetic word problems and the other described strategies that children tend to develop for solving word problems using concrete modeling and counting strategies leading to remembered facts (Knapp & Peterson, 1995). In many cases, once teachers started thinking about their students' understanding of the four basic operations in terms of the frameworks, they began to make different decisions about how to instruct (Franke et al., 2001).

Knapp and Peterson (1995) reported on patterns of CGI usage as a follow-up on the 1989 Carpenter et al. study. They sought to determine if changes in instructional practice and beliefs seen in the original study endured[5]. Knapp and Peterson conducted phone interviews with half of the original CGI participants. They found that three or four years after participating in the workshops, teachers fell into one of three patterns of use. They either (1) saw CGI conceptually and had leveraged their learnings from the workshops to make it the main component of mathematics teaching, (2) became divorced from CGI, seeing it as a set of procedures to be used

---

[5] Note that those teachers who were in the control group in Carpenter et al.'s original 1986 group, participated in the CGI workshop the following year, so all 40 teachers from the original study had participated in training prior to this follow-up study.

as a supplement to traditional practices, or (3) fell away from their original level of CGI usage, even though their beliefs about effective mathematics instruction suggested CGI principles.

Teachers who used CGI as the main component of their mathematics teaching continued to develop their teaching practice after the conclusion of the CGI seminars. They focused on students' development of conceptual understanding; believed elementary teachers needed substantial mathematics understanding to teach well; and allowed student-developed strategies. They used collaborative groups and encouraged interdependence of students. Likewise, teachers who saw CGI as supplementary shared characteristics. They focused on procedural competence, thought pedagogical knowledge in the absence of deep content knowledge was enough to teach mathematics to young children, and demonstrated mathematical procedures for children. Their students worked alone to get the right answers. Teachers who had fallen away from using CGI had incongruencies between what they said they believed and what they did in their classes. Those who saw CGI as the mainstay of their instruction proved that the CGI PD could have an effect on teachers' overall classroom practice in mathematics.

Franke, Carpenter, Fennema, Ansell, and Behrend (1998) provided case studies of three CGI teachers that showed patterns of CGI usage with similarities to those investigated by Knapp and Petersen (1995). Franke et al. used these cases to explain and provide examples of *self-sustaining, generative change* in teaching practice following PD. *Self-sustaining* change is making an instructional change, like allowing multiple solution strategies or having students discuss strategies, then, seeing that students learned from or become more engaged because of the changed instruction, subsequently deciding to maintain that change. *Generative* change occurs when a teacher not only realizes that an instructional change is working, but also strives to understand why the change works, what is different in student thinking, and how instruction might further build on this. Consequently, the teacher makes connections between instructional practice and student learning that can form the basis for future thinking and learning. As explained by Franke and colleagues, “self-sustaining, generative change…frequently entails teachers making changes in their basic epistemological perspectives, their knowledge of what it means to learn, as well as their conceptions of classroom practice” (p. 67). The qualities of the three case study teachers that aligned with whether they demonstrated, self-sustaining or generative change were leveraged in another follow-up study for CGI.

As the ESP researchers followed up with their project participants after the conclusion of the professional development, CGI researchers also followed up with their 1990-1993 seminar participants to investigate whether teachers continued to use the CGI principles. Four years after the teachers’ participation in the CGI workshops concluded, Franke, Carpenter, Levi, and Fennema (2001) interviewed and observed 22 teachers’ patterns of change to determine if there was generative learning and what factors set the generative learners apart from the other CGI teachers. This study used a rating scale for “Levels of Engagement with Children’s Mathematical Thinking” (Franke et al., 2001) to measure the teachers’ level of *generative growth.* Raters used levels 1, 2, 3, 4A, and 4B. Teachers above level 2 all valued children’s thinking as a central tenant of their teaching. Teachers above level 3 used more specificity when describing their children’s mathematical thinking. Teachers at level 4B had generative growth. They viewed children’s thinking as central; possessed detailed knowledge about children’s thinking; discussed frameworks for characterizing the development of children’s thinking; perceived themselves as creating and elaborating on their own knowledge about children’s thinking; and sought colleagues who had knowledge about children’s thinking. Ten of 22 teachers in the study showed generative growth.

They had not just maintained their learning from the previous CGI workshops but had continued to expand and grow their own knowledge about student thinking.

In both CGI and ESP, the PD in which teachers took part led to changes in teacher knowledge and/or beliefs (Boston, 2013; Fennema et al., 1996; Franke et al., 2001). The change in knowledge and beliefs was associated with changes in classroom practice from traditional to ambitious (Boston, 2013; Carpenter et al., 1989). In CGI, it was a focus on children's thinking in the PD that became embedded in teachers' practice. For Boston and Smith "the task-centric approach allow(ed) for gradual, sustained growth along a continuum of task selection and implementation" (Boston & Smith, 2011, p. 974).

2.3 Coaching of teachers can support ambitious instruction

When compared to the literature surrounding ambitious instruction, the literature around coaching of teachers, and especially coaching of mathematics teachers, is not as rich. Although books about coaching and training of coaches are available (e.g., Confer, 2006; Hull, Balka, & Miles, 2009; Morse, 2009; West & Staub, 2003 etc.,) until recently, relatively few empirical studies examined the role of the mathematics coach (Chval et al., 2010; Cobb & Jackson, 2011; Murray, Ma, & Mazur, 2009) or the impacts of coaching on teacher practice and student learning (Matsumura et al., 2010; Neufeld & Roper, 2003; Polly, 2012). As Campbell and Malkus put it, "interest in mathematics specialist-coaches has outpaced not only research studies of their impact, but also clarity in terms of their expected baseline knowledge and professional expertise" (Campbell & Malkus, 2014, p. 215). For this reason, the literature review of coaching expands beyond the limits of coaching of *mathematics* teachers to coaching of teachers in more general terms. First, this portion of the review of literature will examine some of the roots of coaching. Next, the chapter follows the evolution of peer coaching. The chapter then examines potential definitions of coaching as well as other coaching types or models before shifting to focus primarily on content-focused coaching. With regard to mathematics coaching, the literature review will examine the coach's development, and roles and responsibilities, as well as the effectiveness of mathematics coaching, as measured in various studies. This portion of the literature review concludes by relating coaching to ambitious instruction.

Perhaps because of the relative dearth of empirical studies about coaching's effect on teaching practice and student achievement, "the evidence that coaching is an effective strategy for improving instruction and learning remains relatively weak" (Matsumura et al., 2010, p. 36). The review of literature around the impact of coaching on instruction and student learning revealed inconsistent findings. Campbell (2012) concurs with this, stating that "research frequently offers contradictory results" and offering that "one reason for the discrepancy may lie in the differing expectations for these specialists/coaches" (p. 157). Gibbons and Cobb (2016) discuss that even when expectations for coaching are consistent, there are considerable differences among the activities of the coaches in a given setting. The differing expectations and coaching activities are not a recent development. Even the earliest proponents of classroom coaching, Bruce Joyce and Beverly Showers, altered their expectations for coaches and the accompanying model of classroom coaching over the course of their writings. With all its inconsistencies, coaching is still a "promising alternative to traditional models of professional development" (Kraft, Blazar, & Hogan, 2018, p. 547).

2.3.1 Roots of coaching of teachers

The idea of coaching teachers evolved from its introduction within the work of Joyce and Showers. In 1980, Joyce and Showers published an article concerning teacher training in which they present ideas about how and why combinations of five training components–theory, demonstration, practice, feedback, and classroom application–are effective at allowing teachers to either fine tune existing teaching strategies or adopt new strategies. For these authors, coaching occurred to aid the classroom application component of the training. Joyce and Showers (1980) initially define this type of coaching as *coaching for application*. "Coaching for application involves helping teachers analyze the content to be taught and the approach to be taken and making very specific plans to help the students adapt to the new teaching approach" (p. 384). In the 1980 writing, Joyce and Showers name numerous individuals in education-related roles who might serve as coaches; most of whom are knowledgeable others. However, the authors are attracted to the idea that coaching by peers might prove a convenient way to allow teacher change to occur.

Roots of coaching teachers are also found in publications surrounding CGI. During the project, CGI staff and mentor teachers made observations and had informal interactions with participating teachers. The "type of support varied depending on the mentor and the teacher, but included observing in the teacher's classroom and discussing the children's thinking, planning lessons together, and assessing children together" (Franke et al., 1998, p. 71). This is similar to classroom coaching minus the aspect of co-planning classroom lessons with the supportive other. There is evidence in the CGI publications that this quasi-coaching made a difference in some teachers' practice. In fact many high-level, generative teachers from the CGI project cited the support associated with the project as essential to their generative growth and continued level of engagement with children's thinking (Franke et al., 2001). Regarding CGI project teachers who were not generative in their growth, Knapp and Petersen (1995) wondered whether "more scheduled opportunities for 'coaching' and interaction with both researchers and other teachers over the school year might have helped teachers enlarge their interpretation of CGI" (p. 62).

As he did with ambitious teaching, Lee Shulman influenced the development of coaching as a professional development tool. In particular, Shulman (1998) conceptualized the connection between theory and practice in the teaching profession. Shulman noted at the time of his writing that cognitive scientists were thinking about the apprenticeship model for teacher training, as proffered by Dewey in 1904. "Dewey had espoused…that only theoretical learning *situated in practice* would be rich and meaningful" (p. 524). Shulman's research team posed that instead of pairing the trainee with a practicing professional *after* being immersed in the theory of the practice, educators should consider a "cognitive internship" in which the trainee's field experience connects to theory currently being learned to more readily allow for connection and application. Shulman called this "situated intellectual work" because it "embed(s) the learning in the social context of practice" (p. 524). If one envisions applying the notion to in-service teachers involved in professional development instead of limiting it to pre-service teachers in training, this writing can be considered as a precursor of cognitive-coaching.

While there are multiple threads of literature forming possible foundations for classroom coaching, the first of these lines, from Joyce and Showers, provided the initial progression of research studies and publications around coaching. This initial line of research explored the idea of peer coaching in the classroom and connected it to what was then called staff development.

2.3.2 Peer coaching

Joyce and Shower's 1980 publication introduced the idea of coaching, but with their 1982 publication, which was a review of existing literature on teacher training, Joyce and Showers embrace peer coaching, advocating that schools use teams of teachers to provide reciprocal assistance to and support for one another in the classroom. In peer coaching, teacher teams attend training and study a new teaching technique or method; They plan and practice the method with each other, watching demonstrations and working out points of confusion. The teams then take turns watching each other try the new method or technique in their respective classrooms with students. Afterwards, the teachers provide feedback and constructive criticism to each other. This repeats until the teachers develop proficiency with the newly learned teaching strategy. Joyce and Showers defend this method of coaching being paired with in-service training by writing "Coaching without the study of theory, the observation of demonstrations, and opportunities for practice with feedback will, in fact, accomplish very little" (Joyce & Showers, 1982, p. 5).

In 1984, Showers further developed the idea of peer coaching within a study investigating whether (1) teachers can be trained to coach their peers in the classroom application of new teaching strategies; (2) teachers who are coached by peers transfer training at a greater rate than uncoached teachers (following identical initial training); and (3) students of peer-coached teachers perform better on specified tasks than students of uncoached teachers. The study involved 21 teachers and six peer coaches. Findings from the study included: (1) Peer coaches were trained in a relatively brief period to provide follow-up training to other teachers; (2) Peer coaching increased the transfer of training rate for coached compared to uncoached teachers; (3) Students of coached teachers performed better on a concept attainment measure than students of uncoached teachers. However, there were some caveats to these findings. Firstly, the peer coaches in this study had previously been trained in the new teaching technique and had taken a coaching training course prior to doing any coaching. Thus, the peer coaches were actually more knowledgeable than the teachers whom they coached and were not co-equal partners as peer coaching suggests. Secondly, “peer coaches varied considerably in the extent to which they analyzed appropriate use of newly-learned strategies within curriculum areas” (p. 18) with two of the peer coaches providing reinforcement to their coachees to the point where the reinforcing feedback conflicted with the coaches’ “ability to provide accurate feedback on teacher performance following a lesson” (p. 18). With only six coaches in the study and considerable variability, the conclusion that peer coaching is a reliable method for increasing transfer of training may not be valid. Lastly, some teachers in the study chose not to fully participate in the study. This impacted transfer of training to the classroom for those teacher/coach pairs. While Showers’ study (1984) held promise for effective content coaching, it did not show that peer coaching is a model that will improve teacher transfer of training or student performance.

In 1996, Showers and Joyce again claimed to have confirmed their hypothesis about training followed by coaching resulting in more transfer than training alone. They wrote, “teachers who had a coaching relationship…practiced new skills and strategies more frequently and applied them more appropriately than did their counterparts who worked alone” (p. 2). Not only will *training* alone result in less transfer, Showers and Joyce (1996) also claimed that *coaching* alone did not aid in student learning when they write, “There is no evidence that simply organizing peer coaching…will affect students' learning environments.” (p. 1). In this writing, Showers and Joyce recommended coaching first be done with someone who has more expertise, like a consultant or an outside expert. Following that, coaching could be peer-to-peer. Within this 1996 iteration of their coaching model, Showers and Joyce recommend coaching take place weekly and that it be comprised of co-planning, observation, and time thinking about impacts on student learning *without* verbal feedback. The change to exclude verbal feedback was recommended because coaching seemed too much like an evaluation to participating teachers

The same research team reiterated some of the 1996 findings in their 2002 writing, stating that even when a series of "high-powered" (Joyce & Showers, 2002, p. 77) PD sessions possessed the elements Joyce and Showers had deemed to be important (e.g., theory explanations, demonstrations, practice), the effect size was minimal to none, but with the addition of coaching, the effect size was 1.42. Joyce and Showers further added that if new skill or knowledge is to be put to use in the classroom, teachers need coaching, but in this iteration of their on-going endorsement of coaching, the team does not advocate for coaching by trainers, relative experts, or more knowledgeable others. This time, the authors advocate for coaching by co-equal peers, providing the rationale that "coaching by trainers will give the same effects, but is not practical in most settings" (p. 77). In fact, Joyce and Showers theorized that 95% of teachers would transfer their learning to the classroom, if peer coaching was used. Other changes in the Joyce and Showers coaching model appeared in the 2002 writing. For one, the person teaching was now the coach, and the person observing was the coachee. Additionally, peer coaching now consisted mainly of co-planning for lessons. Any conversations or feedback after the lesson is taught were no longer in the Joyce and Showers model for peer coaching. While their stance on who should serve as a coach and the model of coaching they advocated had morphed over the decades of their writing and researching, Joyce and Showers' claim that training programs with a coaching component help teachers better transfer new knowledge and skills to their classroom was a constant.

Kohler, Criley, Shearer, and Good (1997) also conducted a small study on "peer" coaching of four elementary teachers by one common coach. The study found coached teachers more likely to try new techniques and strategies in their classrooms. However, the coach in this experiment, who was also the second author, had 35 years of experience teaching elementary school plus three years of experience in coaching and using the instructional technique under examination (integrated instructional approach or IIA). Because the coach was a relative expert and not a peer to the teachers in the study, the study is *not* truly one of peer coaching.

Conversely, Murray, Ma, and Mazur (2009) undertook an empirical study of peer coaching in which the participants *did* have mutual expertise. The team used an experimental group of nine teachers and a control group of five teachers in a pre- and post-test design to study changes in mathematics or science knowledge of students in peer partners' classes. Peer partners were to work collaboratively, observing one another and providing support and feedback, in an effort to implement what had been learned in a 1- or 2-week summer institute around middle school (grades 7-9) mathematics or science teaching. Quantitative results on sample items drawn from the *Programme for International Student Assessment* (PISA) 2000 and 2003 show no statistically significant difference between pre- and post-test scores for students of treatment group teachers. Qualitatively, researchers found that post-conference sessions between peers lasted an average of 13 minutes and involved an average of 12 topics. The conversations were superficial, lacking any degree of depth, and did not involve constructive criticism. Every conversation was positive "without a single negative comment made" (p. 207). The feedback peer coaches provided to one another during the post-conference was descriptive of the taught lesson. The conversations were not analytical or reflective. "Neither did the observers ask any question that would effectively motivate reflection or analysis" (p. 207).

Perhaps Murray, Ma, and Mazur (2009) contributed to the current trend regarding peer coaching wherein coaching by co-equals has fallen out of favor and been supplanted with coaching by more knowledgeable others (Cobb & Jackson, 2011; Krupa & Confrey, 2012; Polly, 2012). Since Joyce and Showers' 2002 publication, researchers have not embraced the idea of coaching being the process of co-planning and observing without any feedback. In fact, the provision of feedback by the coach has become a critical component of most coaching models (Gibbons & Cobb, 2016; Kraft et al., 2018; Russell, Correnti, Stein, Hannan, & Bill, 2017; West & Staub, 2003). The literature review now examines some of the definitions and types of coaching that evolved from Joyce and Showers' original conception.

2.3.3 What is coaching?

In the time period between the introduction of coaching in the 1980s and the present, multiple coaching types and definitions of coaching developed. Coaching has grown in its use; sometimes connected to outside-the-classroom PD and sometimes used on its own (Kraft et al., 2018). For some schools, districts, or research initiatives, this form of PD has evolved to the point where the role of the coach has become an "important and pivotal resource" (Chval et al., 2010, p. 194). Although "teacher coaching has emerged as a promising alternative to traditional models of professional development" (Kraft et al., 2018, p. 547), results from coaching-related studies vary, perhaps because there are many different definitions (Campbell, 2012; Kraft et al., 2018).

The earliest explanation of coaching is from Joyce and Showers (1982). They name five major functions: provision of companionship; giving of technical feedback; analysis of application; adaptation to students; and personal facilitation. Neufeld and Roper (2003) provide another, more comprehensive explanation, including school leaders in the description. They say

> The term coaching includes activities related to developing the organizational capacity of whole schools (such as increasing leadership for instructional reform). It includes helping

> principals and teachers reallocate their resources and improve their use of data in the service of improving instruction. And it includes activities directly related to improving instruction (such as one-on-one observation and feedback). (p. 4)

Other publications provide explanations specific to *mathematics* coaching. Foster and Noyce, M.D. (2004) provide an explanation of mathematics coaching from the Mathematics Assessment Collaborative (MAC), a consortium of school districts near Silicon Valley for the purpose of improving mathematics instruction via the examination of student work within professional development. Their definition provides an allocation of coaches' time.

> Mathematics coaches are accomplished teachers with records of leadership and strong understanding of mathematics content who are released from teaching duties to work with other teachers…The coaches spend 70% of their time supporting other teachers in the classroom and the remainder either offering professional development to groups of teachers or participating in further professional development of their own. (p. 373)

Hull, Balka, and Miles' *Guide to Mathematics Coaching* (2009) defines a mathematics coach as "an individual who is well-versed in mathematics content and pedagogy and who works directly with classroom teachers to improve students' learning of mathematics" (p. 8). The publication states that coaches have to see the "big picture of mathematics teaching and learning" (p. 5). Coaches "improve the whole by improving component parts" (p. 5) and might be considered change agents. Coaches have many interconnected responsibilities and possess knowledge about content and teaching but also have the social skills to work well with other adults.

Chval, et al. (2010) identify four main components of a mathematics coach's role: supporter of teachers, supporter of students, supporter of school-at-large, and learner, and they draw from a portion of Virginia's adopted description of mathematics specialists as

> teacher leaders with strong preparation and background in mathematics content, instructional strategies, and school leadership. Based in elementary and middle schools, mathematics specialists are excellent teachers who are released from full-time classroom responsibilities so they can support the professional growth of colleagues, promoting enhanced mathematics instruction and student learning throughout their schools. They are responsible for strengthening classroom teachers' understanding of mathematics content, and helping teachers develop more effective mathematics teaching practices that allow all students to reach high standards as well as sharing research addressing how students learn mathematics. (p. 192)

Campbell and Malkus (2014) also call upon Virginia's job description of a mathematics specialist, saying it is all-encompassing and includes co-planning, co-teaching, and debriefing as a part of the role, but the role also includes items as diverse as working with administrators to provide leadership for the mathematics program to interpretation of high-stakes assessment results. Campbell and Malkus sum up the role of specialist-coach as a "collegial mentor who helps foster and then works to sustain a practice-based professional community" (p. 214).

While Gibbons and Cobb's explanation goes beyond mathematics coaching, the researchers focus their definition on one form of coaching, content-focused coaching, and differentiate it from other forms by stating that content-focused coaches "(a) are more knowledgeable partners who have developed relatively accomplished instructional practices (Neufeld & Roper, 2003; Poglinco et al., 2003; West & Staub, 2003) and (b) aim to support teachers' development of ambitious instructional practices in a particular discipline" (p. 239) by "provid[ing] teachers with ongoing, job-embedded support for improving the quality of their instruction and their students' learning" (p. 255).

A recent all-encompassing definition of coaching comes from Kraft, Blazar, and Hogan's (2018) meta-analysis of coaching. They "define coaching programs broadly as all in-service PD programs where coaches or peers observe teachers' instruction and provide feedback to help them improve" (p. 548), but acknowledge that across the spectrum of the 60 studies included in the meta-analysis, there was not a common definition. Sometimes coaching was defined as a partnering of peers, but more often "coaches are thought to be experts in their fields, who model research-based practices and work with teachers to incorporate these practices into their own classrooms (Sailors & Shanklin, 2010)" (p. 551). No matter which definition of coaching was considered, "this is a demanding role, and a role that the profession does not understand and is only beginning to examine" (Campbell & Malkus, 2011).

2.3.4 Types or models of coaching

Over the last decades, many coaching models evolved from peer coaching to the idea of "plac[ing] a highly knowledgeable teacher, who frequently does not have responsibility for the instruction of a classroom of students, in a school in order to advance instructional and programmatic change" (Campbell & Malkus, 2011). Even with the shift towards more knowledgeable others serving as coaches, as of 2015, no clear cut evidence on the effectiveness of one model over others had emerged. According to the National Mathematics Advisory Panel (U.S. Department of Education, 2008), "the Panel found no high-quality research showing that the use of any…types of math specialist teachers improves students' learning" (p. xxii). However, the Panel did not cite research refuting the use of coaches. Hull, Balka, and Miles (2009) used the lack of refuting research, writing, "the key question should not be whether coaching works but under what conditions" (p. 2). Six years hence, Blazar and Kraft (2015) wrote, "Despite growing evidence of the benefits of high-quality coaching, questions remain about the efficacy of different types of coaching programs" (p. 544). The literature review now examines models of coaching along with existing evidence of efficacy.

2.3.4.1 Expert coaching

Polly (2012) performed a study in which he acted as an *expert* coach. The author stated his study was motivated by the need for research evidence to support the efficacy of coaching and to provide clarification about coaching models that are effective. The stated purpose was "to examine the types of support elementary school teachers seek from more knowledgeable others and the influence of various types of support on their teaching while attempting to implement standards-based pedagogies" (p. 81).

Polly (2012) recruited two third grade and two fifth grade mathematics teachers for his study. All of the teachers had six or fewer years of teaching experience. The author informed the teachers of standards-based instruction and said he would support their mathematics teaching in any ways they desired, including co-planning, co-teaching, modeling lessons, or providing curricular resources. Upon teacher request, the author performed observations and provided feedback, taught or co-planned and co-taught sample lessons, and provided curricular resources. At the conclusion of the study, Polly had performed 21 to 30 observations per teacher.

Polly (2012) analyzed the beginning two, middle two, and ending two observations of each teacher to gauge teacher progress. He coded for cognitive demand of tasks (Stein & Smith, 1998) and used his own framework to code for the question types. Over the course of the school year, a greater percentage of enacted tasks were either *Procedures without Connections* or *Procedures with Connections* and a smaller percentage were *Memorization* tasks, which represents a shift towards tasks with greater cognitive demand. All four teachers shifted towards tasks with greater cognitive demand from beginning to middle to ending observations. Additionally, "all four teachers asked more higher-level questions as the year progressed. Specifically, teachers posed more and higher-level questions towards the end of a lesson as students were sharing their work on mathematical tasks" (p. 88). However, no teacher asked any questions of the highest level on the authors' hierarchy until the last observation period when one of the four teachers had 6% of her questions at the highest level. From this small study, Polly concluded that expert coaching has potential because the four teachers taking part in the study shifted to posing higher-level tasks and questions in one year with support.

2.3.4.2 Content coaching

Neufeld and Roper (2003) defined *content* coaches as those who focus on improvements in instruction in a content area by working at the classroom and school level. Furthermore,

> Content coaches do not have a scripted role. They must understand the instructional reform they are helping teachers implement, they must be skillful in working with adult learners who may be skeptical about–or threatened by–the reforms, and they must know how to adapt their coaching methods to the knowledge and skill of the teachers. (p. 3)

Neufeld and Roper also recommended that content coaches establish a non-evaluative environment for teachers, hold small-group PD sessions for their coachees, assist teachers in transferring knowledge attained in outside-the-classroom PD, and "help teachers develop leadership skills with which they can support the work of their colleagues" (p. 9).

Krupa and Confrey (2012) performed a case study about instructional coaching in high school mathematics. They did not classify the coaching type for their study. Because the coaches in the study were relative experts assisting in continuous improvement in mathematics teaching, this study is grouped with content coaching. The model for coaching used in this project shared similarities to another from Loucks-Horsley, Stiles, Mundry. Love, and Hewson (2012) which had five elements: "(1) teachers focusing on learning or improvement; (2) a climate of trust, collegiality, and continuous growth…; (3) coaches are well-prepared, with in-depth content knowledge; (4) mechanisms for observing practice and providing feedback…; and (5) opportunities for interaction" (p. 165).

The Krupa and Confrey (2012) case study grew from the North Carolina Integrated Mathematics Project (NCIM), a two and a half-year project designed to support teachers in rural high schools in implementing a standards-based integrated mathematics curriculum. In this project, teachers attended a summer institute to learn about both content and pedagogy related to using the reformed curricular program. The project created a network for rural teachers which held follow-up conferences and hosted a website to help teachers overcome some of the so-called "challenges of isolation" (p. 162). The project also instituted instructional coaches who were experienced teachers and who made monthly visits to each school. The coaches were relative experts in the use of the standards-based curriculum the teachers were using. They arranged the site visits, observed, and reflected with teachers to meet the individual needs of each teacher.

Examination of coaches' documentation revealed that teachers needed support for many elements of teaching, including content knowledge, planning, questioning, and formative assessment. Coaches' activities with teachers fell into four broad categories: (1) curriculum and content assistance; (2) planning, enactment, and reflection; (3) assessment, feedback, and grading; and (4) professional community interactions (e.g., website use). Interviews revealed that teachers felt most helped by "planning, observing the coaches model teaching, getting access to technology and support in using it, and receiving feedback following an observation" (Krupa & Confrey, 2012, p. 167). This case study provided evidence of the effect professional development plus coaching can have on teachers' knowledge and practice. In particular, the case study teacher used traditional methods even after the first summer workshop. Her coach worked with her to provide resources, plan, provide access to classrooms using ambitious methods, and give feedback. Over the life of the project, the teacher changed her practice and, in fact, became willing to model teach and mentor new colleagues, but the transition was gradual and required working with the instructional coach. The study bolsters Krupa and Confrey's statement: "Research has shown convincingly that teachers are not likely to change their instructional practices solely by attending isolated professional developments, and that ongoing support can help teachers implement the ideas presented in these professional developments" (p. 161).

2.3.4.3 Content-focused coaching

Now, the literature review turns its attention to *content-focused* coaching. Expert coaching, content coaching, and content-focused coaching have some similarities. All involve a variety of possible activities (Hull et al., 2009; Morse, 2009; Polly, 2012; West & Staub, 2003). All employ more knowledgeable others in a coaching role. However, there also exist differences among these coaching types. For example, Polly (2012) allowed the four teachers in his expert coaching study to prompt him regarding their instructional needs and wants. While content-focused coaches differentiate their coaching based on each teacher's background and the coach-coachee relationship, the content-focused coach does not wait for the coachee to prompt the coaching process, nor is the process driven solely by the coachee's desires (West & Staub, 2003).

Content-focused coaching was developed by the Institute for Learning at the University of Pittsburgh (Matsumura et al., 2010). According to West and Staub (2003), content-focused coaches want to improve student opportunities to learn by giving teachers opportunities to improve practice. "under the guidance of skilled mentors" (p. xiv). West and Staub were specific about the goals, roles, and responsibilities of content-focused coaching in their book, *Content-Focused Coaching: Transforming Mathematics Lessons*. The authors recommended that coaching be specifically aimed at what should be taught by teachers and learned by students. Goals for coaching included (a) the design of lessons so students learn something that is a part of the core learning in the content area; (b) the creation of professional habits of mind along with communicative relationships with colleagues; and (c) the development or refinement of the teacher's pedagogical content knowledge.

Gibbons and Cobb (2016) say the "intent of content-focused coaching…is to provide teachers with ongoing, job-embedded support for improving the quality of their instruction and their students' learning" (p. 255). They differentiate content-focused coaching from other forms by stating that content-focused coaches "(a) are more knowledgeable partners who have developed relatively accomplished instructional practices (Neufeld & Roper, 2003; Poglinco et al., 2003; West & Staub, 2003) and (b) aim to support teachers' development of ambitious instructional practices in a particular discipline" (p. 239).

2.3.4.3.1 Research in literacy

The first longitudinal, group-randomized study of content-focused coaching (CFC) with significant positive results for student learning was performed with elementary literacy coaches and teachers. Matsumura, Garnier, Correnti, Junker, and Bickel (2010) and Matsumura, Garnier, and Spybrook (2013) report on the study's findings. The study involved the fourth and fifth grades at 29 schools: 15 treatment schools and 14 control schools. Data sources for the three-year study were numerous. They included (1) multiple sources of feedback and information directly from teachers; (2) a measure of the quality of classroom text discussions via the IQA (Matsumura et al., 2008; Resnick et al., 2006); and (3) student test scores from the Texas Assessment of Knowledge and Skills (TAKS) and Degree of Reading Power Assessment.

The coach-trainees in the study spent time learning about the pedagogy of best practices in reading instruction, learned the skills of coaching teachers, and developed coaching expertise from fellows at the Institute for Learning (IFL). While there was professional development for the coaches provided by IFL, there was no accompanying PD for the participating teachers. The research team attempted to have participating coaches avoid non-coaching tasks by having principals attend professional development with coaches from their buildings. Throughout the study, coaches were expected to hold weekly grade-level meetings and have a monthly coaching cycle (plan, teach, reflect) with each teacher. (Matsumura et al., 2010).

The study was not without complications. Matsumura et al. (2010, 2013) created a second cohort of new teachers between years 1 and 2 of the three-year study because teacher turn-over in participating schools became problematic. Also, at the end of the second year, few teachers reported full participation in monthly coaching activities or weekly grade-level team meetings.

Despite the complications, the study bore positive results. Many teachers did participate in a level of activity close to the desired level. Teachers participated in coaching between four and six times a year and met in grade-levels once a month or more. This was significantly higher than in the comparison schools and impacted instruction as the research team hoped. “By the end of the second year of the program, text discussions in the CFC schools were more interactive and rigorous than in the comparison schools” (Matsumura et al., 2013, p. 44). This led to significant increases in student achievement as measured on the TAKS (Matsumura et al., 2013).

2.3.4.3.2 Research in mathematics

An impactful study of content-focused coaching in elementary mathematics took place between 2005 and 2008. It measured the effect of elementary mathematics coaches on student achievement (Campbell, 2012; Campbell & Malkus, 2011), on teacher beliefs (Campbell, 2012), and on coaches' content knowledge, mathematical knowledge for teaching, and beliefs about teaching and learning (Campbell & Malkus, 2014). Like the Matsumura et al. (2010, 2013) study in literacy, this was a three-year, randomized control study in elementary schools. Also, like the Matsumura et al. study, this one provided PD for coaches but not for teachers. Instead of pairs of schools, this study used 12 triples (36 schools) to provide two different cohorts of coaches in experimental schools with one control school per triple. Coaching began in 2005. Coaches in ten of the original 12 schools continued for the three years of the study as cohort 1. Cohort 2 coaches received training in 2006 and began coaching the following school year, so at the conclusion of the study, coaching in cohort 2 schools had been in place only one year.

Student achievement in grades 3-5 at the cohort 1 experimental schools increased over time. Although test scores in comparison schools were higher than in control schools after the first year, significant increases in student achievement were not yet evident. The increases occurred as coaching became enculturated in the schools (i.e., coaches gained experience, school staff learned to work together). Increases became significant and were maintained in grades 4 and 5 after year 2 and became significant in grade 3 after year 3 (Campbell & Malkus, 2011).. For the cohort 2 schools, where coaches had been in their role for one year, having a mathematics coach did not significantly impact mathematics achievement scores (Campbell, 2012). "The pragmatic implication of this finding is the caution that a coach's positive effect on student achievement develops over time as a knowledgeable coach and the instructional and administrative staffs in the assigned school learn and work together" (Campbell & Malkus, 2011). This finding dovetails with a Chval et al. (2010) study showing that a coach's identity develops over the first year of being in the role.

Campbell (2012) analyzed teachers' beliefs about mathematics teaching in the same set of schools. To measure changes in beliefs, researchers used a survey on which teachers rated 30 items from strongly disagree to agree. The items reflected perspectives about mathematics teaching ranging from "Traditional" views to what the authors called "Making Sense" views, labeled as such because they aligned with views espoused in *Making Sense* (Hiebert et al., 1997). Teachers' views did not change with regard to either perspective of mathematics teaching, unless they were "highly engaged with the specialist" (p. 156). "The beliefs of teachers who were highly engaged with a specialist changed significantly, shifting away from the Traditional perspective toward a Making Sense perspective" (p. 156).

Also using the same set of schools and coaches, Campbell and Malkus (2014) examined changes in coaches themselves as they trained for and transitioned to their coaching roles. Prior to placement, coaches in the Campbell and Malkus (Campbell, 2012; Campbell & Malkus, 2011, 2014) study took a leadership course and five mathematics content courses. Courses emphasized a "Making Sense" approach, meaning participants were to reason, solve problems, participate in discourse, work collaboratively, and make connections among solution strategies and among mathematics concepts. The coaches learned content, but they experienced learning in a way that focused on ambitious teaching. Coaches took another leadership course with an emphasis on coaching during their first year of placement.

Before and after participating in the first set of courses, coaches took (1) a test of content knowledge; (2) a paper-pencil MKT assessment; and (3) a beliefs survey. After each year of coaching, the coaches again took the beliefs survey and the MKT assessment. Results showed a statistically significant increase on the content knowledge assessment in both cohorts of coaches. Over time, results also showed significant increases on MKT scores for both cohorts. Furthermore, for both cohorts beliefs were impacted significantly, becoming less traditional and more aligned with a "Making Sense" perspective. The combination of courses taken and experience in the role of a coach seemed to impact content knowledge, MKT, and beliefs.

The effects on coaches resulting from this study suggest PD beginning in advance of coaching responsibilities and continuing during coaching impacts coaches' knowledge and beliefs regarding mathematics and mathematics teaching and learning. The student achievement and teacher beliefs resulting from this study, along with Matsumura et al.'s (2010, 2013) results, suggest (1) the greater the time mathematics coaches spend on coaching, the greater their impact on student achievement, with the caveat that the first year bears no real fruit because of the transition to the new role and responsibilities; and (2) affecting teachers' professional growth requires a coach establishing, developing, and maintaining relationships with coached teachers.

Also in the area of mathematics, Gibbons and Cobb (2016) performed a case study with an instructional coach in middle-school mathematics. The case study was extracted from the MIST study on how to support mathematics teachers' in becoming more ambitious and equitable in their classroom practice (Cobb & Jackson, 2011). Gibbons and Cobb's case study uncovered five *coaching* practices the case study coach engaged in during planning that impacted her content-focused coaching. Those coaching practices were: "(a) identifying long-term goals for teachers' development, (b) assessing teachers' current instructional practices, (c) locating teachers' current instructional practices on general trajectories of teachers' development, (d) identifying next steps for teachers' development, and (e) designing activities to support teachers' learning" (p. 246) which involved modeling, co-teaching, and observing and debriefing with teachers. Gibbons and Cobb also named knowledge of ambitious teaching of mathematics and of the general trajectories mentioned in (b) and (c) as two areas of knowledge impacting the coach's planning.

2.3.4.4 The TN + IFL Math Coaching Model

The most comprehensive mathematics coaching model published to date grew from a collaboration among IFL and the Learning Research & Development Center (LRDC), both at the University of Pittsburgh, and the Tennessee Department of Education (TN DOE) and was called the TN Math Coaching Initiative (Russell et al., 2019). This model has roots in the content-focused coaching model from IFL but has additional elements which help scale the model for widespread use. Figure 2.3 displays the complete TN + IFL Math Coaching Model. The middle column of the figure shows the *coaching practice framework* which is briefly explained here and discussed in some detail in upcoming sections. The coaching practice framework is comprised of three parts: a *coach development framework*, which will be explained in greater detail in section 1.7.5; a *coaching framework*, which will be explained in greater detail in section 1.7.6; and "an ethos of *continuous improvement* that informs how coaches are trained to use disciplined inquiry cycles to adaptively integrate the coaching model into their diverse local contexts" (p. 5).

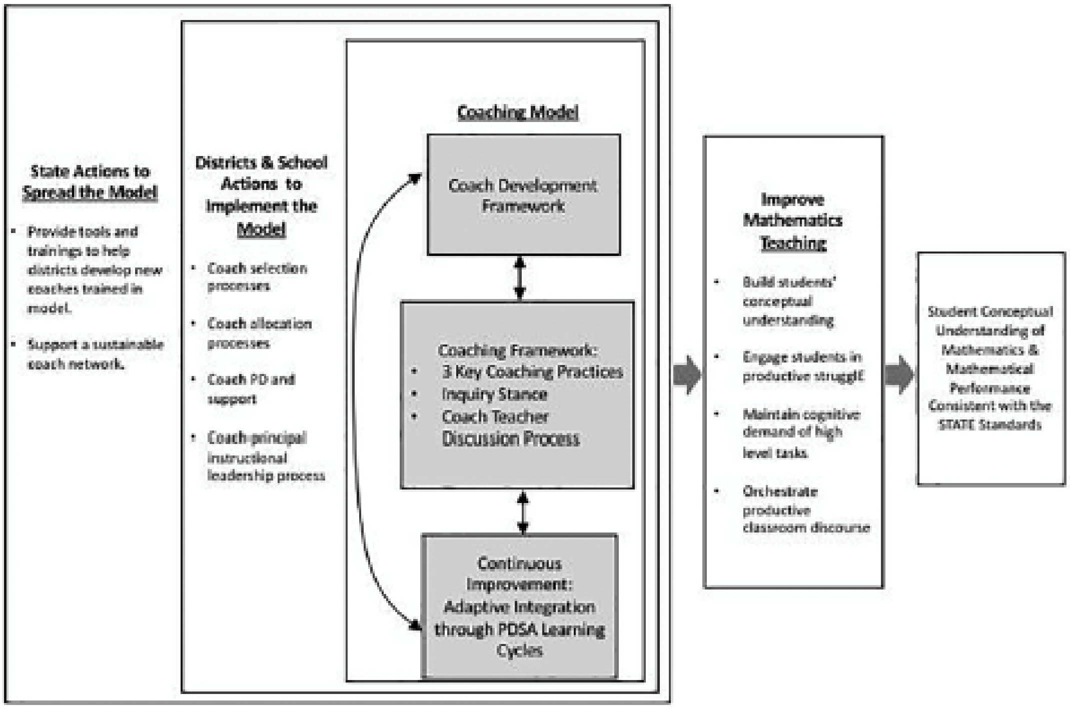


Figure 2.3 The TN + IFL Math Coaching Model (Russell et al., 2019, p. 6)

Improvement science and the Plan-Do-Study-Act (PDSA) learning cycle of adaptive integration from Bryk, Gomez, Grunow and LeMahieu (2015) influenced the continuous improvement portion of the coaching practice framework for the Russell et al. (2019) study. Using continuous improvement science in a large scale study made this model nearly unique in the mathematics coaching literature. (Cobb and Jackson (2011) applied a continuous improvement model in developing their theory of action for improving mathematics instruction at scale in the MIST project. While their theory of action goes beyond coaching, mathematics coaching *was* one of the five key components in Cobb and Jackson's theory of action.) The TN Math Coaching Project began developing, testing, and refining this TN + IFL Math Coaching Model in 2014 with 32 mathematics coaches from 21 school districts. The coaches in the project each committed to work closely with two partner teachers. The partner teachers all taught mathematics in grades 3 through 8 for a total of 63 partner teachers. Aside from working intensively with two partner teachers, the coaches also provided documentation of all coaching activities, completed periodic surveys, and attended monthly webinars and three two-day meetings each year.

Within the continuous improvement process, researchers gathered data from coaches and teachers about what coaches were doing during the coaching cycle. Knowing what the coaches had been trained to do, the researchers analyzed the effectiveness of the coaching process by examining planning documentation and videos from planning sessions, teaching lessons, and debriefing conversations. From these data sources, the researchers were quickly alerted to challenges and issues coaches encountered in attempting to implement the coaching framework. Then, the coaching framework within the larger, overall coaching model was adjusted to accommodate the findings. As Russell et al. (2019) put it, "researcher-driven inquiry cycles attended to variation in implementation of the model and sought to identify adaptations that were associated with positive coaching and/or teaching outcomes, which could become part of the model's design" (p. 10-11). The process of examining data and fine-tuning the coaching model became a feedback loop with researchers and then coaches participating. Having a clear coaching model at the outset was important for scalability, but the model needed to be adapted once the relationships between coaching and its outcomes were established via data, so "the essence of the model" (p. 29) could surface.

Regardless of the coaching model, many of the studies showing positive results, including the Russell et al. (2019) study, had coaches partake in professional development. Even though the teachers being coached were not necessarily getting PD outside the classroom, outside training was having an influence on teacher practice via in-the-classroom coaching. With that in mind, the literature review now turns to examining research and publications concerning the transition to becoming a coach and the roles and responsibilities inherent therein.

2.3.5 The transition to coaching

Joyce and Showers (1980) say coaches need to be collaborative and non-judgmental. Loucks-Horsley et al. (2012) say there needs to exist a climate of trust, long-term commitment to coaching, and support from administration for coaching to succeed. Hull, Balka, and Miles (2009) say coaches have to see the "big picture of mathematics teaching and learning" and they "improve the whole by improving component parts" (p. 5). Furthermore, Hull, Balka, and Miles say coaches might be considered change agents who have lots of interconnected responsibilities. They have to possess knowledge about content and teaching, but also need to have social skills to work well with other adults. Campbell (2012) says

> coaches must also learn how to support teachers while questioning them; how to frame a common goal across differing instructional philosophies while trying to build community within and across grade-level teams; how to facilitate positive discussion advancing mathematical knowledge while addressing teachers' limited understandings; and how to navigate the organizational and cultural factors that exist in schools. (p. 150)

Otherwise, added Campbell and Malkus (2014) "as has been found within peer coaching…, interactions within a coaching dyad may be positive but lacking in the level of analysis and reflection needed to advance or change a teacher's understandings or classroom practices" (p. 221-222). Lastly, Confer (2006) says "Our goal as math coaches is not to add a little spice, salt, or pepper to the stew of mathematics instruction, but instead to alter the menu entirely" (p. 2). Taken together, these writings make the job of the mathematics coach seem insurmountable, especially because, as Chval et al. (2010) said, one of the issues inherent to coaching is that coaches don't step out of the classroom ready to coach.

As far back as Showers' (1984), one finds mention of training for coaches, however; in the peer coaching model, Showers stated that "coaches can be trained in a relatively brief period" (p. 48). This stands in stark contrast to the on-going and in-depth training provided within a number of the coaching studies reviewed in this chapter (Campbell & Malkus, 2011; Foster & Noyce M.D., 2004; Krupa & Confrey, 2012; Russell et al., 2019), wherein coaches attend weeklong or weeks-long training, sometimes every summer, with follow-up meetings or other PD throughout the school year. In fact, Campbell (2012) wrote that when coaches did not participate in professional development, there was minimal impact on student achievement.

Aside from training, transitioning to coaching sometimes requires others with whom the coach will work to learn and adjust. Neufeld and Roper (2003) provided multiple recommendations regarding coaches' development when writing, "principals and coaches [should] understand the 'big picture' of the reform in which they are engaged and the reasons that undergird the changes" (p. 11). Russell et al. (2019) echoed that others with whom the coach works need to support coaching. Their coaching model involved not only district and school actions but state actions to propagate the model. See Figure 2.3 for the TN + IFL Coaching Model showing district and state actions. Campbell (2012) wrote that coaches may not be as effective during their first year because there are many challenges associated with the transition from teacher to coach, including, but not limited to, enacting their training. Campbell said there exist "additional, distinct abilities…to be effective coaches of other teachers" (p. 150). Further, there is a shift in identity that involves changing "from being viewed by others and by oneself as an expert (as an expert teacher) to being viewed as a novice (a novice specialist or coach)" (p. 150). Chval et al. (2010)

added information about how the coach shifts identities over the first year of tenure in the same way as changes occur when a teacher transitions to an administrator.

Through meetings, conversations, and surveys, Chval et al. (2010) identified four components of a mathematics coach's identity that are different from a mathematics teacher's identity but are related to the coach's new and developing role within the school's culture. The first is *coach as supporter of teachers*. The new coaches anticipated this would be the biggest part of their new job. The second recognized identity was *coach as supporter of students*. Coaches had to let go of their teacher identity to develop their coaching identity. The next identity was *coach as learner*. While the coaches taking part in this study recognized that a necessary part of their job was continued growth, coaches sometimes felt concern about how other teachers might view it. The last identity Chval et al. identified was *coach as supporter of the school-at-large*. Initially, coaches envisioned this as creating a school-wide vision of mathematics instruction. However, this identity involved various duties in support of the school that were unanticipated (e.g., making copies, cleaning the cafeteria) and did not seem related to the role of a mathematics coach. All in all, initially perceived identities were different from what the identities became in reality over the course of the first year in the position.

2.3.6 Roles and responsibilities of coaches

Campbell and Malkus (2011) provided a *raison d'etre* for the mathematics coach. "Elementary mathematics coaches are placed in schools to construct leadership roles and to provide on-site, collaborative professional development addressing mathematical content, pedagogy, and curriculum in an effort to enhance instruction and improve student achievement" (p. 430). Aligned with this reason for their existence, Neufeld and Roper (2003) described a number of broad roles and responsibilities for coaches at both the classroom and school level. Russell et al. (2019) extended the reach of support structures beyond the school and district level to the state level in their TN + IFL Coaching Model.

Much like the definitions for coaching, the roles and responsibilities of coaches and those who support the work of coaches at school, district, and state levels are varied and numerous. In fact, Gibbons and Cobb (2016) cite this as a possible reason why results from empirical studies are inconsistent. However, in attempting to specify impactful roles and responsibilities, the authors uncovered "three potentially productive activities that coaches might enact one-on-one with teachers in their classrooms: (a) co-teaching, (b) modeling, and (c) debriefing" (p. 240). *Co-teaching,* said Gibbons and Cobb (2016), can help teachers because instructional practices are impacted when a teacher witnesses what a coach does with the teacher's own students or when a teacher works with a coach to plan the actions each will take during the lesson implementation and then witnesses what the students do in response. *Modeling* can be especially fruitful for encouraging ambitious instructional practices when the coach calls attention to actions he or she takes during the model lesson and to students' responses or reactions. *Debriefing* a teacher-taught lesson with respect to challenges encountered can also be a fruitful coaching practice. By working with more knowledgeable others, teachers generate potential solutions to problems they might encounter during instruction, and the coach can provide specific pointers on a given ambitious teaching practice. It should be noted, according to Gibbons and Cobb, being an exemplary teacher is insufficient, although necessary, for developing the coaching expertise needed to successfully engage in modeling, co-teaching, and debriefing.

Both Kennedy (2016), in her review of research on effective PD, and West and Staub (2003), in their content-focused coaching book, wrote about the behaviors of and expectations for effective coaches. The identified behaviors were not all synonymous with those from Gibbons and Cobb (2016). Kennedy emphasized collaborative planning and goal-setting. Like Kennedy (2016), West and Staub (2003) focused on goal-setting and collaborative planning, but they added task analysis as an expectation for effective coaching, writing that the coach

> must know the mathematics in depth and be able to show teachers how to set specific learning goals for a lesson, devise or select powerful tasks, analyze the knowledge–correct and 'misconceived'–that children are likely to bring to the tasks, and plan instructional conversations that are contingent on student responses. (p. xiv)

To this end, "coach and teacher collaboratively plan, enact, and reflect on specific lessons, acting as resources for each other" (p. 2).

Goal setting, using rich tasks, and collaborative planning are all coaches' roles within the TN + IFL Math Coaching Model, but there is even more to the coach's role in this model. The TN Math Coaching Project identified three main elements of their *coaching framework* encompassing the roles and responsibilities of coaches during the coaching process. (See the middle box in Figure 2.3.) These three elements of the *coaching framework* are: a set of *3 Key Coaching Practices;* the *coach-teacher discussion process;* and an *inquiry stance.* The *3 key coaching practices* are: "(1) deep and specific discussions of the instructional triangle, (2) establishing mathematics and pedagogical goals, and (3) evidence-based feedback" (Russell et al., 2019, p. 5). Each of the *3 key coaching practices* enters the coaching process at particular points during the *coach-teacher discussion process*, which is an updated, more sophisticated version of the plan, enact, reflect coaching cycle. The *coach-teacher discussion process* is shown in Appendix B. Throughout the discussion process, the coach maintains an *inquiry stance*. According to Russell et al. (2019), taking an inquiry stance involves using noticings and wonderings as opposed to giving direct instruction. “The inquiry stance stems from…the need for active teacher participation in meaning making around shifts in practice” (p. 6).

2.3.6.1 The coaching cycle

Enacting the coaching cycle of plan, enact, and debrief with teachers is the most visible part of a coach’s roles and responsibilities. West and Staub (2003) were not the only writers to discuss the coaching cycle. Confer (2006), in *The Math Coach Field Guide*, reified the detailed co-planning championed by West and Staub saying that the teacher brings expertise regarding the individual students that the coach probably does not have, but the coach may bring content or teaching knowledge the teacher can learn. Hull, Balka, and Miles (2009) used one chapter in their *Guide to Mathematics Coaching* to discuss planning and co-teaching and one chapter to discuss analysis and reflection in the coaching cycle.

Campbell and Malkus (2014) also discussed a plan, teach, reflect cycle for mathematics coaches. They created a conceptual model, shown as Figure 2.4, illustrating the nature of this work as well as the learning occurring for both coach and coachee throughout the process. This model shows similarities to the cycle described by West and Staub (2003) in the cyclical nature of the planning, teaching, and reflecting occurring on the right side of the figure as well as in the elements of each phase in the coaching cycle.

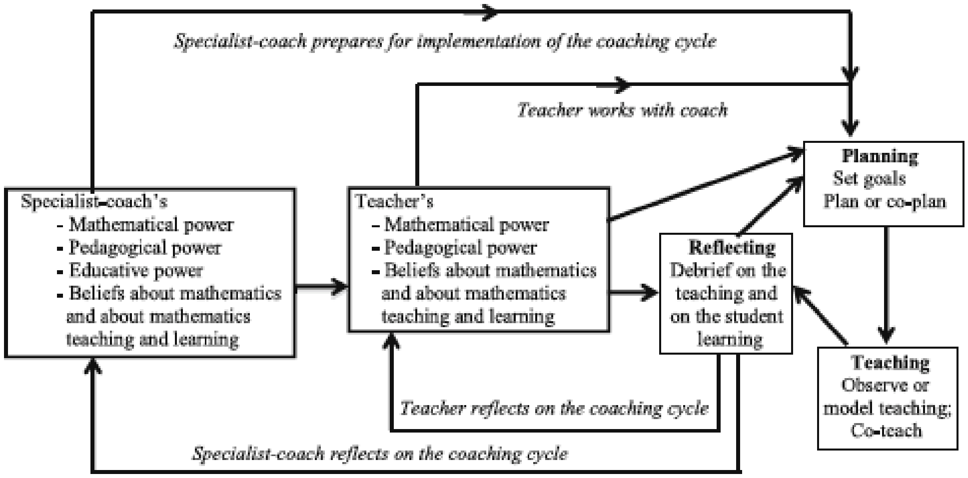


Figure 2.4 Model of coach and teacher co-learning in coaching cycle (Campbell & Malkus, 2014, p. 217)

Other similarities to West and Stuab elucidated in the Campbell and Malkus study are in the background the coach and teacher bring to the process and the learning that occurs throughout the process. The coach brings educative power–"the additional, accessible knowledge that a teacher educator needs in order to support teachers as they solve or address mathematical and pedagogical problems" (p. 216). Both coach and teacher bring mathematical and pedagogical expertise and their current beliefs about mathematics and mathematics teaching and learning to the coaching process. Beliefs are influenced by prior experiences and drawn upon during planning and teaching. Reflecting upon the lesson then contributes to changes in the beliefs and in the content and pedagogical knowledge each brings to the subsequent coaching cycle.

Russell et al.'s (2019) version of the coaching cycle, called the *coach-teacher discussion process,* is one of the three elements of the *coaching framework*, along with *inquiry stance* and *3 key coaching practices*. See Figure 2.3 for a depiction of the overall TN + IFL Coaching Model, where the *coaching framework* is in the middle of the figure. The *coach-teacher discussion process,* as depicted in Appendix B, begins with coach and teacher identifying the mathematics learning goal–which is one of the *3 key coaching practices*–and selecting a cognitively demanding task for use during the lesson. Each then completes the task and identifies solution strategies students might use to solve the task. Next, there is a pre-observation planning conference, wherein deep and specific conversations of the instructional triangle (D. K. Cohen, Raudenbush, & Ball, 2003) are to occur. The deep and specific conversations are another of the *3 key coaching practices*. Then, the lesson observation takes place followed by the post-observation feedback conference. During the feedback conference, the final of the *3 key coaching practices*, using evidence-based feedback, is facilitated via the use of an evidence based collection tool (Bill, personal communication). Evidence-based feedback is paramount during the lesson analysis within the feedback conference. Ever-present during the *coach-teacher discussion process,* is an *inquiry stance* of noticing and wondering on the part of the coach.

West and Staub (2003) encouraged two kinds of coaching moves be used during the coaching cycle's co-planning or reflecting phases, the first of which foreshadows the *inquiry stance* that umbrellas the *coach-teacher discussion process* of the TN Math Coaching Project. "Moves that invite the teacher to verbalize his or her perceptions, thoughts, plans, deliberations, and arguments" (p. 46) have the potential to encourage teachers to construct their own meaning regarding potential changes in professional practice much like the inquiry stance. The other move encouraged by West and Staub relates to some unforeseen results of the Russell el al. (2019) study. "Moves through which the coach provides direct assistance relevant to the planning and implementation of the lesson" (West & Staub, 2003, p. 46), in other words, being explicit with teachers about what they should or should not do in the lesson, is not aligned with the inquiry stance advocated by Russell et al. However, results from Russell et al. showed no difference in growth patterns for teachers who experienced increased explicitness from their coach and the overall teacher group, demonstrating that "under the right conditions, explicitness may be efficacious; as such, it called for a reconsideration of the principle of taking an inquiry stance with teachers, a component of the coaching model's design" (p. 22).

In addition to studying the effect of explicitness, Russell et al. (2019) used the iterative cycles of coaching and continuous improvement (Bryk et al., 2015) to study coaching press. Coaching press served as an indicator that coaches were having deep and specific conversations with teachers about the instructional triangle–one of the *3 key coaching practices*–and were pushing the teachers to construct their own ideas about effective mathematics instruction. In contrast to the results for an inquiry stance, results *did* show that increased press of the teacher by the coach during the coaching cycle resulted in positive change for teachers.

As is apparent in the Russell et al. (2019) study, among others, "this approach to professional development is complex and requires considerable thought as well as ingenuity in order to take the core idea and create an effective coaching model" (Neufeld & Roper, 2003, p. 19). One complexity, sensed by Showers over 30 years ago, was that "coaching is not a simple additive that can be tacked on to the school with a 'business as usual' attitude, but rather represents a change in the conduct of business" (1985, p. 26). This complexity is being addressed as researchers such as Campbell and Malkus, Russell and her team, and Cobb and his team examine systems so coaching can be effective, sustained, and taken to scale, resulting in a "change in the conduct of business" for schools. The literature review now turns to examining studies showing the effectiveness of coaching, including a meta-analysis of 60 coaching studies.

2.3.7 Effectiveness of coaching

Much has changed since Joyce and Showers first laid claim to coaching's effectiveness in the mid-1980s. Studies of coaching have used increasingly sophisticated methodology and analysis (e.g., Matsumura et al., 2013; Russell et al., 2019). They have made use of reliable and valid tools (e.g., IQA from Boston, MKT assessment from the LMT group). They have expanded their scope and reach (e.g., MIST project, TN + IFL Math Coaching Project), and despite the fact that the number of empirical studies of coaching still pales in comparison to the number of empirical studies of ambitious instruction, they have become more numerous, to the point where meta-analyses have now been performed with coaching studies.

One such meta-analysis was published by Kraft, Blazar, and Hogan (2018), who "limited" their research to 60 coaching studies, most of which were randomized control studies. Using meta-analysis methods allowed Kraft et al. to increase the statistical power afforded to the combined results, examine pooled results from a variety of coaching models, and compare the characteristics of models that may be related to effectiveness. Results from the Kraft et al. study showed "pooled effect sizes of 0.49 standard deviations (SD) on instruction" (p. 547) with an Interquartile Range (IQR) from 0.17 to 0.92 SD, and pooled effect sizes of "0.18 SD on [student] achievement" (p. 547) with an IQR of 0.03 to 0.24 SD. However, the effect sizes increased by 0.31 SD for instruction and 0.12 SD for students achievement when coaching was paired with group training. In general, effects on student achievement were less than effects on classroom instruction in this group of studies. Additionally, effect sizes for larger studies were associated with smaller effect sizes for both changes in teacher instructional practice and student achievement. The majority of studies in the Kraft et al. (2018) meta-analysis focused on literacy, as did Matsumura et al.'s (2010, 2013) study showing positive effects of coaching on both classroom instruction and students achievement. Only two of the studies included in this meta-analysis focused on mathematics. Regarding those studies, students' mathematics achievement showed a very small positive effect of 0.050.

While Matsumura et al. (2010, 2013) studied content-focused coaching's effect in literacy, Campbell (2012) and Campbell and Malkus (Campbell & Malkus, 2011, 2014) reported on content-focused coaching in the mathematics classroom with positive outcomes for student achievement, teacher beliefs, coach beliefs, and coach content knowledge and MKT. Campbell and Malkus' results showed effects of mathematics coaching at each juncture of Desimone's (2009) proposed conceptual framework for studying professional development. As such, Campbell drew on the Desimone framework to create her own model showing how the mathematics coach impacts teachers' learning in PD settings to ultimately impact teachers and students. Campbell's model, as shown in Appendix C, depicts the coach's interactions with teachers in three forms of professional development–one-on-one or grade-level coaching, school-based PD, and larger-scale PD. All these forms of coach/teacher interaction around professional learning impact the teacher knowledge and beliefs, which, according to Desimone, impact instruction and feed back to impact professional development. The instruction ultimately impacts student achievement, which also feeds back to impact teacher knowledge and beliefs. Lastly, the increased student achievement feeds back to impact instruction. Thus, effective coaching has consequences beyond the teacher, as shown by Campbell and Malkus.

This literature review has made it apparent that (1) there are many models for coaching in schools, and those models have morphed and become more sophisticated; (2) there are many definitions or explanations of coaching that have also changed over time, but a consensus seems to have formed that the coaching is "instructional experts work(ing) with teachers to discuss classroom practice in a way that is (a) individualized…; (b) intensive…; (c) sustained…; (d) context specific…; and (e) focused" (Kraft et al., 2018, p. 553); (3) the training, roles, and responsibilities allocated to coaches are also varied and can be intense; but despite, or perhaps because of, the evolving nature of coaching, (4) some studies have shown coaching as an effective form of PD to change teacher knowledge, beliefs, and instructional practices and even increase student achievement. The literature review concludes this section on coaching by examining literature making the connection between coaching and ambitious mathematics instruction more salient.

2.3.8 Coaching in relation to ambitious instruction

Aside from already-reviewed literature making the direct connection between coaching and ambitious instruction, (e.g., Campbell & Malkus, 2014; Gibbons & Cobb, 2016; Kraft et al., 2018; Matsumura et al., 2013), Cobb and Jackson (2011) provided a "comprehensive, empirically grounded theory of action for improving the quality of mathematics teaching at scale" (p. 6) that involves mathematics coaching. This theory of action for district-level instructional improvement was created and refined over a years-long time period in association with the previously-mentioned MIST project. The theory of action explicated five key components for improving mathematics education at scale by making classroom instruction more ambitious. One of these five key components was "mathematics coaches' practices in providing job-embedded support for teachers' learning" (p. 9). Other key components involved teacher networks and the practices of both school and district leaders in support of ambitious instruction. These components, including that of mathematics coaches' practices, all support the first key component for Cobb and Jackson's theory of action: “a coherent instructional system for supporting mathematics teachers’ development of ambitious teaching practices” (p. 10-11).

In explaining their ideas about the coherent instructional system, Cobb and Jackson (2011) provided seven associated goals. The first of these was a set of explicit goals for students. The second was a vision of high-quality instruction with specific practices. The goals and vision “should drive the design of the remaining elements of the instructional system” (p. 13), and the vision is one place where Cobb and Jackson’s theory of action connected ambitious instruction to coaching. The authors recommended the district vision encompass an agreed-upon but “relatively small set of high-leverage instructional practices that are learnable in the context of high-quality professional development” (p. 13). These high-leverage or ambitious instructional practices should influence instructional leadership and PD that includes district-wide professional development, coaching, and Professional Learning Communities.

Cobb and Jackson’s (2011) main point in making coaches a part of the five key components was to be clear about the fact that those who have already developed sound instructional practices, otherwise known as the mathematics coaches, should work with teachers in the classroom to develop the high-leverage, ambitious teaching practices that are a part of the coherent instructional system (the first of the five key components in this theory of action). Cobb and Jackson go on to state that participating in a coaching cycle of co-plan, enact, and analyze, akin to the cycle previously discussed in this chapter, could be productive in supporting ambitious teaching and that coaching activities may be even more productive if paired with district professional development. Accordingly, the district PD should support ambitious instruction, and the messages delivered to teachers via the professional development in all its forms should be reinforced in word and deed by district leaders and school instructional leadership. All in the organization–teachers, coaches, school leaders, district leaders–must take a learning perspective and adjust their practice to further ambitious teaching in the name of instructional improvement, say Cobb and Jackson.

Cobb and Jackson's (2011) theory of action may be the most explicit one connecting coaching to ambitious mathematics instruction, but for at least a decade before their publication, others have implicitly connected ambitious instruction to forms of PD inherent in coaching. For example, Smith (2001), in her book encouraging *Practice-Based Professional Development*, considered designing professional development while keeping in mind the "cycle of teachers' work and the nature of the activities in which teachers engage as they move through the cycle" (p. 8). This cycle is aligned with the cycle of coaching activities as it begins with planning, continues with the actual teaching, and concludes with a reflection. Later, she reiterated the call to provide help for teachers that focuses on their everyday activities and stated, "This type of assistance can be provided by supportive ventures that focus directly on an individual teacher's practice, such as coteaching, coaching, assistance with planning, and reflection on actual lessons" (p. 42). Thus, Smith provides assistance for those attempting to help teachers instruct in what was labeled a *standards-based* way and what has become known as *ambitious*.

In addition to Smith (2001), Borasi and Fonzi (2002) discussed scaffolded field experiences (SFEs) as a PD technique. They defined SFEs as "opportunities for participating teachers to experiment with instructional innovation while receiving support" (p. 83) and claimed SFEs can help teachers learn to use effective teaching strategies, writing that until teachers have an opportunity to try a new teaching practice in their classrooms, they cannot know what it is really like, even if they have viewed videos or seen the practice modeled by others. The SFEs that encourage effective, ambitious instruction are aligned with the practice of coaching because the teachers being coached are experimenting while receiving support.

Borasi and Fonzi’s (2002) publication has commonalities to Grossman et al.’s (2009) study, the purpose of which was to develop a framework describing and analyzing the teaching of professional practice within education programs for relational professions. Grossman et al. set out to discover how practice was taught by gathering data from different preparation programs. The team identified three components common across the programs: *decomposition*, *approximation*, and *representation* of practice. Decomposition is the breaking down of practice into its component parts. Decomposition of practice is a part of Grossman et al.’s framework because the actual practice of these professionals is so complex that learners new to the practice need to have the chance to distinguish the components that make it up. Representations are made of the ways the profession shows the practice to the learners along with the pieces of the practice that are made visible. “Approximations of practice refer to opportunities…to engage in practices that are more or less proximal to the practices of a profession” (p. 2058).

Coaching encompasses Grossman et al.’s (2009) framework because it (1) allows teachers and coaches to work together in discussing the components that make up ambitious instruction as well as those components pertinent to lessons being planned. This encompasses representation and decomposition. (2) Coaching allows teachers to try out the components of their teaching practice with extensive guidance from the coach during planning and during one phase of the lesson (e.g., the launch, the share-and-discuss). Thus, ambitious instruction can gradually take hold for the teacher. This is akin to Grossman et al.’s decomposition and approximation. Also in coaching, the teacher attempts to put all the pieces of ambitious instruction together when the coach and teacher implement the plan in the classroom. This demonstrates approximation and representation. Although Grossman et al., Smith (2001), and Borasi and Fonzi (2002) are not as explicit as Cobb and Jackson (2011) about how ambitious mathematics instruction and coaching are connected, it is clear from all these writers that coaching is helpful in changing teachers' practice from traditional to ambitious.

2.4 Conclusions

There exists a preponderance of information, in the form of texts, research and even internet sites, related to mathematics teachers' development of ambitious instructional practices. Some of this information relates teacher practice to student achievement (see section 2.2); some relates PD to teacher practice (see section 2.3); some relates teacher knowledge to teacher practice or students achievement (see section 2.4); and some concerns quantifying ambitious instructional practice (see section 2.5). There is not as much information, especially in the form of empirical research, concerning mathematics coaching, but, as evidenced in this literature review, there is an ever-increasing amount. In fact, some research has connected content-focused coaching in mathematics and ambitious instruction. For example, Gibbons and Cobb's (2016) study outlined two aspects of coaching knowledge content-focused coaches need to bring to their coaching practice and five key practices that coaches should use to support the teachers with whom they work in developing more ambitious instructional practices. Cobb and Jackson's (2011) "empirically grounded theory of action for improving the quality of mathematics teaching at scale" (p. 6) involved mathematics coaches as a form of job-embedded PD and on-going support for teachers within a "coherent system of supports for ambitious instruction" (p. 6).

In addition to the inside-the-classroom PD and on-going support from coaches, Cobb and Jackson's (2011) "coherent system of supports" included outside-the-classroom district PD around the desired instructional practices. Kraft et al. (2018) found studies pairing coaching with outside-the-classroom PD, in the area of literacy, were more effective in changing teacher practice and increasing student achievement. Putnam and Borko (2000) have also asserted that teachers should do a portion of their learning in the school or classroom and part outside the school, in other PD settings. Based on the *situative perspective* that states learning is situated, social, and distributed, if teachers confine their learning to coaching, their learning may be limited in its applicability. Putnam and Borko suggested that it may be the combination of approaches that holds the most promise for teacher learning resulting in changes in practice, when they wrote

> If the goal is to help teachers think in new ways,…it may be important to have them experience learning in different settings. The situative perspective helps us see that much of what we do and think is intertwined with the particular contexts in which we act. The classroom is a powerful environment for shaping and constraining how practicing teachers think and act. Many of their patterns of thought and action have become automatic-resistant to reflection or change. Engaging in learning experiences away from this setting may be necessary to help teachers "break set"-to experience things in new ways. (p. 6)

The proposed study is in line with the suggestions of Putnam and Borko (2000) as well as the recommendation in Cobb and Jackson’s (2011) theory of action that mathematics coaches’ practices be included in a comprehensive strategy supporting ambitious mathematics instruction at scale. Additionally, few of the studies reviewed in mathematics education paired outside-the-classroom professional development in mathematics for teachers with inside-the-classroom coaching. Thus, while there are studies showing that professional development and on-going training for *coaches* paired with coaching impacts mathematics teaching (Campbell & Malkus, 2011; Russell et al., 2019), a place is still available within the existing field of research in mathematics teaching and learning in the area of study described in this dissertation pairing outside professional development for teachers with coaching in mathematics.

3.0 Methods

This study seeks to investigate whether and how content-focused coaching affects K-5 mathematics teachers' use of ambitious teaching practices by responding to the following research questions.

- How does proximal, in situ professional development in the form of content-focused coaching paired with outside-the-classroom professional development facilitate a change in math teachers' pedagogical practices from traditional to more ambitious in nature?
    - What is the impact on teachers' *opportunities to learn about* ambitious teaching practices when content-focused coaching is added to professional development?
    - What is the impact on teachers' *use of* ambitious teaching practices when content-focused coaching is added to professional development?

This chapter outlines the approach to inquiry for the proposed study. First, the chapter outlines the study design, including a description of the participant-teachers. Next, information pertaining to and procedures for data collection are shared. Then, the chapter provides information about the instruments used during the study. Lastly, the chapter offers plans for analysis of the data collected.

3.1 Study design

The study described within this chapter is an intervention study with a pre-post design. The change explored in this study is an alteration in teachers' pedagogical practice in K-5 mathematics classrooms from more traditional practices to more ambitious teaching practices (Cobb & Jackson, 2011; Lampert et al., 2011) as set forth in the eight effective teaching practices for mathematics (National Council of Teachers of Mathematics, 2014). To study the potential changes in pedagogical practice, the author engaged in coaching *and* observing one group of five teachers (referred to as the coaching group) and observing only a second group of five teachers (referred to as the comparison group). The samples were drawn from a larger cohort of educators participating in a Title II B Math and Science Partnership (MSP) grant awarded to the Math & Science Collaborative (MSC), where author is employed. Figure 3.1 provides the flow of the study.

As a part of the MSP grant, the MSC delivered 14 days of professional development (PD) to grades K-5 mathematics teachers. The cohort of teachers from which this study's coached and comparison groups were drawn began attending PD during Summer 2017. Before any PD began, all participants took a survey, called the MSP-MSC Survey (University of Pittsburgh, 2016), regarding their current beliefs about teaching and learning mathematics and the Learning Mathematics for Teaching assessment (LMT) (Hill et al., 2004), regarding their current level of mathematics knowledge for teaching. The ten summer days of PD focused on common and specialized content knowledge (Ball et al., 2008). However, the summer PD also implicitly involved pedagogy in a number of ways. For example, the tasks chosen for use in the PD were often tasks that could be used in K-5 classrooms or adapted for use in the elementary grades. Consequently, teachers often raised questions or made comments related to what their students might say or do when provided similar opportunities to learn mathematics. One such activity was a “number talk” (Parrish, 2016) for 83 – 56. Participants noted that very few group members used the traditional US algorithm involving “borrowing” when solving this problem mentally. The teachers questioned its utility as a standard algorithm for students and concluded that instruction should allow for students’ sense making around the operation as opposed to memorizing steps in a solution process.

Furthermore, while the participants did not discuss the connection between these pedagogical practices and the effective teaching practices for mathematics, at the conclusion of every other day, the facilitators of the summer PD asked the following closing question: “What are same ways we learned that (1) helped you make deeper sense of the mathematics than you might have otherwise? (2) helped you think of a different strategy than you might have used before? (3) kept you engaged?” Participants consistently responded that being asked to engage with challenging tasks instead of being told how to solve a problem deepened their sense-making. They often responded that having facilitators question them instead of simply saying if an answer was correct or incorrect helped them think about different strategies or kept them engaged with the mathematics. Thus, although not explicit, the connection to effective Mathematics Teaching Practices was present in the summer portion of the common PD experience. Additionally, discussion engendered by the closing question allowed teachers to make the connection between their own depth of learning in the summer PD and what was possible for their students. Therefore, although the summer experience was meant to deepen teacher content knowledge, it also allowed for opportunities to deepen the pedagogical content knowledge. This set of experienced formed a common foundation for all teachers in the cohort.

Aside from the summer PD, all the teachers in this cohort also participated in four follow-up PD sessions during the 2017-18 school year focused more on pedagogical content knowledge but still involving some specialized content knowledge (Ball et al., 2008). As stated above, a smaller group of ten teachers, drawn from the larger cohort, participated in this study; Five in the coached group, and five in the comparison group. Once the two groups of five were finalized, the author visited each classroom to videotape one mathematics instructional period. Finally, before any coaching began, the author and another coder independently assessed the classroom instruction using four of the Academic Rigor (AR) rubrics from the Instructional Quality Assessment (IQA) (Boston, 2012a) (See Appendix H). The two raters compared and came to consensus on IQA AR ratings for each coached and each comparison teacher.

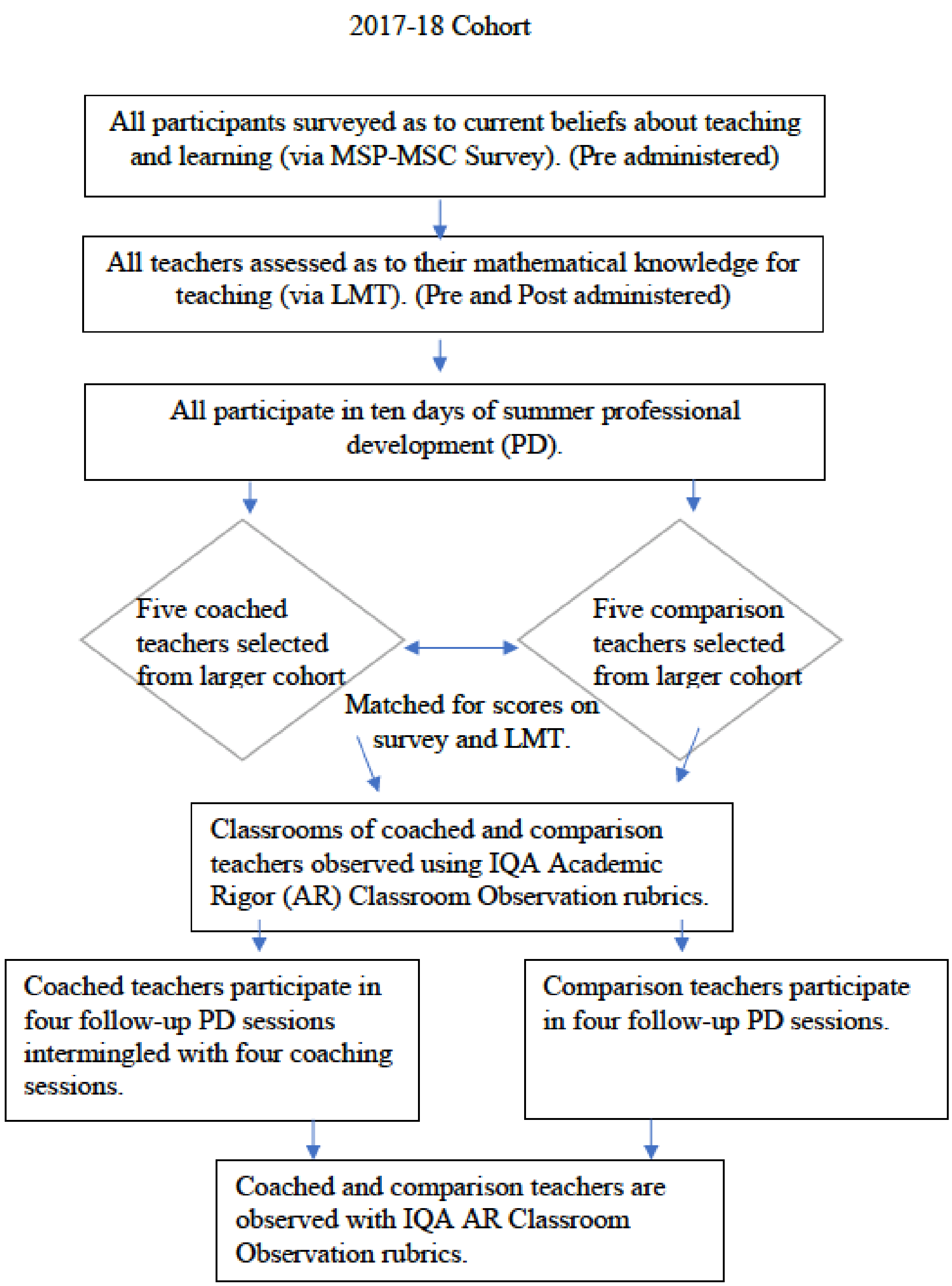


Figure 3.1 Overview of the Study

The figure provides a framework for the study design.

Comparison teachers participated in the four remaining PD sessions during the school year. While the coached teachers participated in the follow-up PD like the comparison teachers, they were also involved in coaching activities intermingled with the PD. After the coaching activities and PD activities were completed for the 2017-18 school year, one additional mathematics class was videotaped. Again, the instruction was assessed by the author and outside evaluator using the same four IQA AR rubrics. Figure 3.2 shows the activities in which the coached and comparison teachers took part as well as the data collection that took place during the 2017-18 school year.

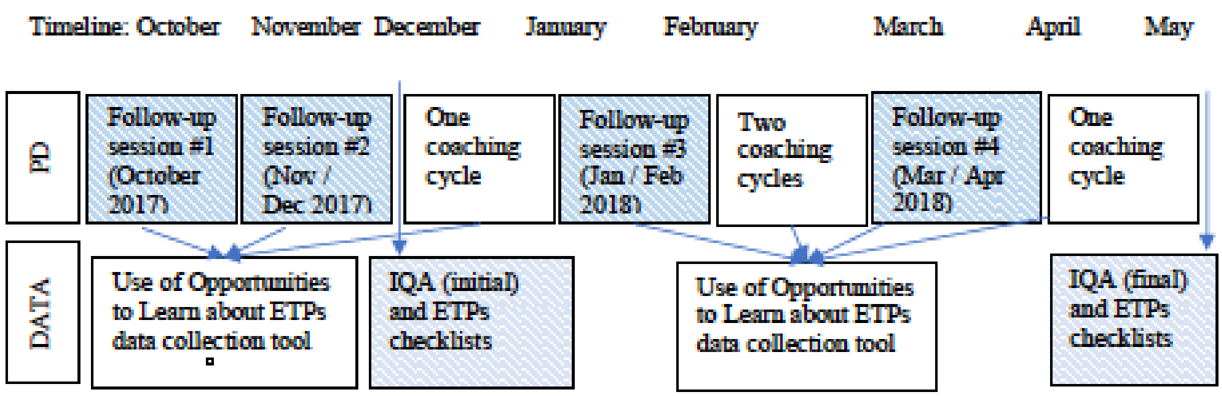


Figure 3.2 Timeline for the study

The figure displays a timeline of professional development and data collection activities for the teachers in this study. The shaded boxes for PD and data collection activities apply to all participants. The unshaded blocks of activities apply only to the coached teachers in the study.

3.1.1 Participants

As previously stated, before the ten days of summer PD, all teachers were administered the LMT assessment (Hill et al., 2004), which provides a measure of teachers knowledge of mathematics for teaching (Ball et al., 2008). Teachers also completed a survey concerning their beliefs about teaching and learning called the MSP-MSC Survey (University of Pittsburgh, 2016).

Scores from the LMT and the survey provided a means for creating two matched groups of five teachers–one group to be coached and one for comparison. Excerpts from the LMT assessment are provided in Appendix D and excerpts from the MSP-MSC Survey are provided in Appendix E.

One teacher in each group had an LMT score more than one standard deviation lower than the mean of the overall group of K-5 teachers receiving professional development. Two teachers in each group had scores within one standard deviation below the mean for the group. One teacher in each group had a score within one standard deviation above the mean, and the remaining teacher in each group had a score more than one standard deviation above the mean of the overall group. The survey scores for each pair of teachers were also within one standard deviation of one another with the exception of the teachers whose LMT scores were more than one standard deviation above the mean. Those scores were within 1.5 standard deviations of one another. While the groups of teachers were intended to be matched, the study did not intend to match teachers one-to-one between coached and the comparison groups, meaning that teachers whose survey and assessment scores matched did not necessarily match for grade level or school district type, even though efforts were made to have teachers from like schools and from the same grades represented in each group.

The teachers participating in the study came from urban-like school districts or suburban districts. Two teachers in the coached group and two teacher in the comparison group taught in urban districts. Three teachers in each group taught in suburban school districts. Two teachers in each group taught grade 1. One teacher in each group taught grade 3, and one teacher in each group taught grade 5. The remaining teacher in the coached group taught grade 2, while the remaining teacher in the comparison group taught grade 4. Coached teachers were not drawn from the same schools as comparison teachers to avoid any contaminating effects of the coaching and to avoid any ill will, since coaching involves additional time and effort on a teacher’s part. In other words, there was no school building with teachers in both the coached and comparison groups. See Appendix F for a summary table of data for teachers in the study.

3.1.2 Coach qualifications

While the coach in this study did not engage in the same initial coach training or the same on-going elements of coaching training as the coaches in Tennessee, from the Russell et al. (2019) publication, the coach for this study did participate in an abbreviated version of that training. Victoria Bill, who oversaw this coach's Supervised Internship (EDUC 3012) at the University of Pittsburgh's IFL provided a shortened version of the Tennessee coach training as well as the opportunity to participate in some initial coaching sessions as an observer and then as a coach-in-training.

The coach for this study did not participate in the same year-long program of studies as the coaches in Campbell and Malkus' (2014) study; however, the coach for this study has received on-going experiences both in graduate programs and in the workplace to enhance her pedagogical content knowledge for teaching mathematics and for leading professional development. Additionally, the coach for this study did engage in some training that overlaps that of the coaches in the Campbell and Malkus studies. For example, the Campbell and Malkus coaches took *Developing Mathematical Ideas* (DMI) courses and facilitator training. The coach for this study participated in DMI courses and also took facilitator training from Dr. Deborah Schifter, and Dr. Virginia Bastable who, along with Dr. Susan Jo Russell are the creators of the DMI PD program and accompanying DMI materials. The coaches in Campbell and Malkus' studies were also trained in use of the Fosnot (2007) instructional materials. While the coach for this study has not received that training, she is familiar with and regularly uses the *Contexts for Learning* (Fosnot, 2007) materials in PD sessions with elementary mathematics teachers.

Likewise, while the coach for this study has not received formal CGI training like the university liaisons in the CGI studies (Franke et al., 1998), she has, however, been steeped in the CGI research and frameworks and has used them in professional development with elementary mathematics teachers. The coach also has more than a decade of previous experience delivering outside-the-classroom professional development. Much of that professional development was based on the tenets of standards-based, ambitious teaching, and much of the professional development was of high quality according to Desimone's (2009) framework. The coach had training in delivering that type of professional development; however, the coach had more limited training or experience specifically geared towards content-focused coaching. Thus, the coach for this study has had some similar and some different experiences compared to coaches in the impactful mathematics coaching studies.

3.1.3 Coaching

Figure 3.3 displays the coaching cycle as implemented for this study. Each coaching cycle consisted of three parts with distinct activities taking place: co-planning, co-teaching, and debriefing. The coaching cycle began with the co-planning or preparation phase of the lesson. Co-planning prepares the coach and teacher for and leads into the co-teaching or implementation phase. Once the lesson has been taught, the coach and teacher debrief on the lesson and any related ideas.

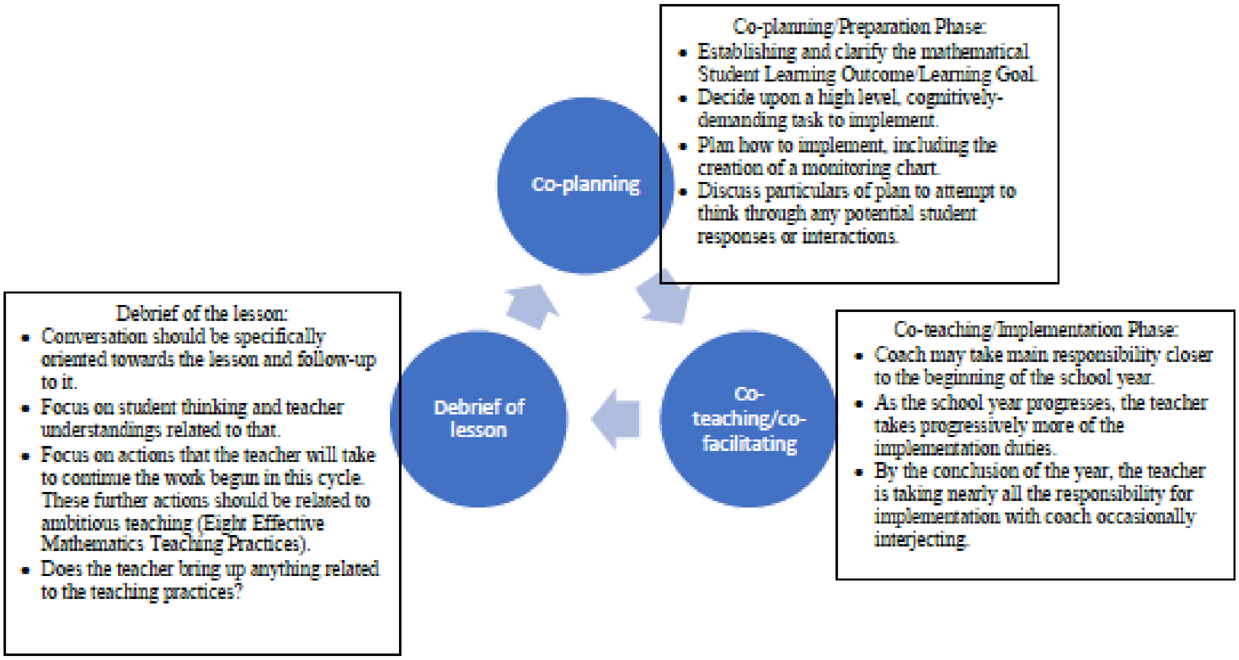


Figure 3.3 Coaching Cycle

3.1.3.1 Co-planning

The first consideration in the preparation phase is the student learning goal for the lesson to be taught. Establishing a clear goal to focus student learning is one of the eight effective Mathematics Teaching Practices. According to *Taking Action: Implementing Effective Mathematics Teaching Practices* "(l)earning goals inform the important decisions teachers make in planning and preparation for instruction, implementing lesson activities, and guiding student learning" (Huinker & Bill, 2017, p. 17). Thus, without a clear learning goal, subsequent planning decisions about which tasks or activities to use, how to launch and facilitate the chosen activities, which purposeful questions to ask, and how else to support students are potentially less impactful. Therefore, the co-planning or preparation phase of the coaching cycle begins before the coach and teacher meet face-to-face with the teacher attempting to craft a learning goal for the upcoming lesson. The pair communicates via electronic means (e.g., phone, email, video chat) to come to consensus on the goal. Because establishing the learning goal is related to each of the other effective Mathematics Teaching Practices, without this target, engaging in ambitious teaching becomes more difficult.

Once the learning goal has been established, the teacher, with some consultation from the coach, chooses the task or activity for the lesson. Because this inquiry and the related coaching activities encourage ambitious teaching of mathematics as embodied in the eight effective teacher practices for mathematics (NCTM, 2014), the chosen task should "promote mathematics reasoning and problem solving and allow multiple entry points and varied solution strategies" (NCTM, 2014, p. 17). While the teacher engaged in such tasks and examined student work related to such tasks during the outside-the-classroom PD, she or he may have had little to no experience implementing such tasks with students. Thus, although learning about cognitive demand of tasks was a part of the follow-up PD sessions, portions of the co-planning phase consistently entailed continuing discussions of the Task Analysis Guide (Stein et al., 2000).

Once the task is selected, teacher and coach separately engage in the task, anticipating both correct and incorrect student solution strategies and crafting assessing and advancing questions related to the anticipated strategies (M. S. Smith & Stein, 2011). A planning template, called the Monitoring Tool (M. S. Smith et al., 2009), records these strategies and questions. The tool helps shift the emphasis of the lesson from teacher actions to student thinking and provides a common protocol from which teacher and coach speak when meeting face-to-face.

After the coach and teacher each create their Monitoring Tool, they meet face-to-face to compare and discuss the strategies and questions they each crafted. At this point, it becomes incumbent upon the coach to push the teacher to think deeply about the students' possible thinking. The coach asks about the details behind what students who use a given strategy might be thinking.

The coach asks about what the teacher anticipates students will say in response to an assessing question she or he has planned to use in association with a given strategy, and then asks, “What if the students don’t say that? What then?” The coach might suggest another strategy–one the teacher did not anticipate–and some questions related to it in order to deepen the teacher’s thinking regarding the task, the student thinking, or even the mathematics of the task. Decisions about what avenues to pursue are made based on a number of different factors, including, but not limited to: (1) the teacher’s expressed interest in a certain effective teaching practice; (2) trouble spots the teacher has previously encountered; or (3) the coach’s determination, based on past coaching episodes, of which effective teaching practice should be pursued.

By anticipating student solution strategies and then comparing, discussing, and delving deeply into these strategies, it becomes more likely that the task, when implemented, will maintain its level of cognitive demand (Stein et al., 1996). Both teacher and coach are likely to routinize the problematics parts of the task or move the emphasis of the task from problem solving into completeness or correctness if they thought in advance about what students might say or do in the process of engaging in the productive struggle (NCTM, 2014). While thorough advanced planning is important for maintaining the level of cognitive demand, it is also critical for other components of ambitious mathematics instruction. For example, learning goals are less likely to be attained if teaching is done spontaneously instead of pre-planned. The instruction may end up being haphazard or disorganized instead of purposefully aimed at student learning (Huinker & Bill, 2017). Additionally, adaptations may not be considered for students whose thinking is less or more sophisticated than most. Thus, the co-planning stage is critical to a coaching episode meant to encourage ambitious mathematics teaching.

3.1.3.2 Co-teaching

Soon after the co-planning process is completed (preferably the same day), the teacher and coach implement the task. During the initial coaching cycle, the coach may take on most of the responsibility for the implementation of the lesson, while the teacher takes on more of the observer's role. The teacher may launch[6] the lesson and then stay at the coach's side during the explore or monitoring and summarize or share and discuss phases (M. S. Smith et al., 2009). As the study progresses, the coach will gradually relinquish more of the responsibilities for lesson implementation to the teacher. The teacher can learn from both observation *and* experience over the next coaching cycles. For example, during the second cycle of coaching, the teacher may launch the task and take on more of the monitoring responsibilities while students are working cooperatively in pairs or small groups. The coach may be at the teacher's side suggesting some of the assessing or advancing questions that have been pre-planned, or the coach may ask an advancing question to the teacher's assessing question to push students further towards the established learning goal.

During the next cycle, the coach may step back even further during the monitoring or explore phase but still consult with the teacher about which student groups to select for sharing and in which order to have them share. Eventually, the coach may step back to only asking a few connecting questions during the share and discuss phase of the lesson. By the conclusion of the

[6] Lessons co-planned and implemented in this study used a launch-explore-summarize model. In the launch, students are drawn into the context of the task, prior knowledge is activated, expectations are established, and the teacher assures that students understand the task and expectations. In the explore phase, students continue to process the problem, mostly in collaborative groups, try solution strategies and refine their thinking. The teacher is circulating during the explore phase so as to monitor student thinking and support students with well-placed questions or comments. In the summarize phase, the class reconvenes into a large group to share, discuss, compare, and contrast the strategies employed by peers in solving the task. The teacher's role becomes pulling the students' ideas together and facilitating the discourse to advance and deepen the understanding of all in the class (Michigan State University College of Natural Sciences, n.d.).

study, the teacher takes all or nearly all the responsibility for implementing the planned task with the students. At this point, the coach becomes more of an observer of the lesson who may occasionally interject a question or comment during the lesson implementation. Grossman et al (2009) used the analogy of learning to kayak on smooth waters which applies here as the coach allows the teacher the opportunity to learn from her and from her implementation of the lesson.

3.1.3.3 Debrief of the lesson

Whether the coach is implementing much of the lesson or taking on the observer's role, it is important for coach and teacher to reflect on the lesson implementation soon after its conclusion. This is a time for the coach to push the teacher to think deeply about the effective Mathematics Teaching Practices that lead to more ambitious instruction. The coach has considered, in advance, the different ways the lesson might play out, much like the teacher and coach jointly consider the implementation of the task. The coach thinks in advance about how to use the lesson in enabling the teacher to more effectively implement future lessons for the advancement of student learning. Doing so most likely entails facets of multiple effective mathematics teaching practices. The coach considers which of these facets to explore with the teacher and which to save for another time. The coach considers how to approach the conversation and connect it to student thinking related to the implemented task. This is all done to further the teacher's progress towards ambitious mathematics teaching.

The debrief allows coach and teacher to deliberately reflect upon the practice of teaching. The coach is the vehicle through which the teacher becomes "aware of what they and their students are doing and how their actions and interactions are affecting students' opportunities to learn" (Huinker & Bill, 2017, p. 251). The coach and teacher jointly consider strengths, weaknesses, and points of potential focus in the future. As cited in *Taking Action* (Huinker & Bill, 2017, p. 251), Hiebert, Morris, Berk, and Jansen (2007) suggest questions to assist in reflecting upon lessons.

- What are students supposed to learn?
- What did students learn?
- How did the teaching help (or not help) students learn?
- How could teaching more effectively help students learn?

These questions assist the teacher in focusing on the students and their learning.

Huinker and Bill (2017) make the point that although the first two questions in Hiebert et al.'s list are likely discussed more often, the last two questions warrant more time and energy in the debrief conversation. These questions focus on how the teacher actions or inactions affect students' learning. "Basing reflections on evidence of teaching actions and student learning helps the teacher form hypotheses about the effects of particular teaching actions on students' learning and to identify ways to improve specific teaching moves" (p. 259) This is the purpose of the coaching debrief. Therefore, focusing on these questions is a potential catalyst to more ambitious teaching practices, whereby students are provided more opportunities to learn.

Although the coach has a teaching practice or two in mind as the focus for each teaching cycle, it is impossible to totally separate the effective mathematics teaching practices from one another. As the teacher and coach work on supporting productive struggle, they must have a task that supports reasoning and problem solving. When working to facilitate meaningful discourse around the task, purposeful questions must be posed and student thinking must be elicited and used to further the productive discussion aimed at an established mathematical goal. However, as the main focus of the cycles shift over the course of the school year, the coach gains formative information about how the effective mathematics teaching practices are developing collectively and individually. Perhaps the teacher shows more facility with choosing tasks that align with the student learning goals, but continues to struggle with questioning strategies. Perhaps the teacher begins to ask purposeful questions but then has trouble using the student responses in a productive way to further the learning. As this information becomes apparent, the coach adjusts the foci of the coaching cycles to bolster any continuing needs on the part of the teacher.

3.2 Data Collection

During the study, data was collected via four sources: the follow-up PD sessions, an initial classroom observation, coaching sessions, and a final classroom observation. Details concerning each of these data sources follows.

3.2.1 Follow-up professional development

Data collection for the study overlapped with the selection of participants for the coaching and comparison groups. All teachers in the larger PD cohort began attending follow-up PD sessions in October, 2017. These follow-up PD sessions provided opportunities for all the teachers in the cohort, including those who were being selected for participation in the study, to learn about the effective mathematics teaching practices. A tool, called the Opportunities to Learn about Effective Teaching Practices Data Collection Tool (OtL-ETP), tracked which of the effective mathematics teaching practices cohort teachers encountered via the follow-up sessions.

Four follow-up sessions occurred throughout the 2017-18 school year; one in October, one in November or December; one in January or February; and the final follow-up session in March or early April 2018. The OtL-ETP tool recorded any effective mathematics teaching practices encountered therein by all cohort teachers, which included all ten of the study's participating teachers. More information about this tool is within the section on Instruments in this chapter, and the tool itself is in Appendix G.

3.2.2 Classroom observations

Following the selection of participants for the study, data collection began with an observation of each participating teacher's classroom. One mathematics instructional period was videotaped in the five classrooms where coaching would occur and in each of the five comparison teacher's classrooms. Beginning in December, 2017 and continuing into early 2018, the author and another trained rater, evaluated the mathematics instruction. The raters used four of the AR rubrics from the IQA tool (Boston, 2012a) Specifically, the AR rubrics are the ones associated with Potential of the Task (AR1), Implementation of the Task (AR2), Student Discussion following Task (AR3), and Rigor of Teachers' Questions (AR-Q). The data gathered via classroom observations served as the starting point for comparisons in the study.

Beginning in late April and continuing until late May, after the conclusion of coaching activities and follow-up PD, data collection concluded with a final observation in each participating teacher's classroom. As with the initial observation, one mathematics class period was videotaped. These classroom episodes were examined and evaluated by the author and the same trained rater as the initial classroom observations. Once again, the raters four AR rubrics – AR1, AR2, AR3, and AR-Q – from the IQA tool.

3.2.3 Coaching

Coaching provided opportunities for the coached teachers to encounter the effective Mathematics Teaching Practices via the co-planning, the implementation of the task, and the debrief of the lesson. The coach used the OtL-ETP tool to record data about these opportunities. Additionally, the coach and coached teacher created qualitative data about each coaching cycle via the creation of Monitoring Tools for the planned lessons. The coach created more qualitative data with planning notes for the co-planning session with the teacher, with planning notes for the debrief, and with notes taken during the co-planning or debrief sessions connected to each coaching episode.

3.3 Instruments

Two instruments: The Opportunities to Learn about Effective Teaching Practices data collection tool and the Instructional Quality Assessment (Boston, 2012a; Boston et al., 2015; Matsumura et al., 2008) provide the bulk of the data related to this study. Notes and planning tools are additional sources of information and evidence related to the study. Another instrument, called the Effective Teaching Practices checklist was intended to be used during the study but was not validated and consequently, data from this tool was not used in the study.

3.3.1 Opportunities to learn about effective teaching practices

The first tool used in the study is called the Opportunities to Learn about Effective Teaching Practices data collection tool (OtL-ETP). The purpose of this tool is to compare and contrast coached and comparison teachers' chances to learn about the eight effective Mathematics Teaching Practices. Data from this tool frames the response to the first research sub-question concerning teachers' opportunities to learn about ambitious teaching practices: *What is the impact on teachers' opportunities to learn about ambitious teaching practices when content-focused coaching is added to professional development?*

Data gathered from the OtL-ETP data collection tool, provide a quantitative measure of how many opportunities to learn about each of the eight effective mathematics teaching practices each teacher experienced during the life of the study. Because each opportunity to learn is assigned a rating of "some" or "extended," the tool also provides some qualitative information concerning the depth or extensiveness of each opportunity to learn about a given effective Mathematics Teaching Practice. Additionally, the tool displays notes, taken by the coach, pertinent to each opportunity to learn about any given effective Mathematics Teaching Practice. The full version of the OtL-ETP tool is in Appendix G. An excerpt from page 1of the tool is shown in Figure 3.4. The tool lists each of the eight effective Mathematics Teaching Practices in abbreviated form in the first column. (e.g. Establish mathematical goals to focus learning is abbreviated as "goals.") The second column rates the learning opportunity provided to the teacher. The last column shows the notes taken in conjunction with the experience.

Opportunities to learn about effective teaching practices data collection tool

Name of Professional Development session (e.g. coaching #x, follow-up #y) Date:

Effective Teaching Practice (ETP)	Opportunity to Learn? (No; Some; or Extended)	Notes about opportunity to learn about ETPs
ETP 1: Goals	Example: SOME	Example: Coach created three possible goals statements related to the topic for the lesson, and together the coach and teacher analyzed each and chose the goal.
ETP 2: Tasks		
ETP 3: Representations		
ETP 4: Discourse		

Figure 3.4 Excerpt from the Opportunities to learn about effective teaching practices data collection tool

This tool was employed in conjunction with the four follow-up professional development sessions and the four coaching sessions for each coached teacher[7]. Following each professional development follow-up session and each coaching cycle, the coach reviewed the presentation notes and reflected upon the discussions among the participants to determine whether or not there was an opportunity to learn about any of the effective mathematics teaching practices. In addition to recording which effective Mathematics Teaching Practices teachers encountered during follow-up professional development sessions and coaching, the OtL-ETP data collection tool documented the depth or intensity of the learning opportunity. Each learning opportunity was assessed as "some opportunity to learn" or "extended opportunity to learn."

[7] As previously mentioned, the follow-up PD was preceded by ten days of summer PD. The summer PD is not included in data set because it focused on content and specialized content knowledge not pedagogical content knowledge via the effective teaching practices for mathematics. However, the author does acknowledge that the summer PD did implicitly involve pedagogy.

An “extended opportunity to learn” is a deep or lengthy encounter with a practice. It may entail multiple connected parts or one in-depth activity or conversation. For example, all the teachers in the study had an extended opportunity to learn about *implementing tasks that promote reasoning and sense-making* during the second follow-up of the school year. During that session, the facilitator introduced the Benchmark Task Grid (Boston, personal communication) (See Appendix L) and facilitated activities to engage participants in rating sample tasks, describing characteristics of the task(s) that make them rate as higher or lower on the grid, and come to consensus on which example tasks were of higher or lower demand and why. Next, the facilitator introduced the Task Analysis Guide (M. S. Smith & Stein, 2011), examines research about task usage from Boston and Wilhelm (2015), resurfaced a number of tasks in which participants had engaged during the summer PD (e.g., *Joey’s Run, Shamrock Smile Mile, Box of Clay;* See Appendices K.5, K.6, and K.7), and compared and contrasted two different place value worksheets to illustrate differing cognitive demand. Then, the participants engaged in *The Hungry Caterpillar Task* and read and discussed the *Case of Ms. Bouchard* (NCTM, 2014). To conclude the day, the facilitator circled back to some research from the QUASAR (Quantitative Understanding: Amplifying Student Achievement and Reasoning) project (Stein & Lane, 1996) to assure teachers they do not have to be perfect when implementing tasks. Teachers’ homework was to find a high-level task and implement it. This opportunity to learn about *implementing tasks that promote reasoning and problem solving* was lengthy and entailed multiple connected parts. Thus, the rating became: “extended opportunity to learn.”

“Some opportunity to learn” entails something shorter in duration or an encounter with less depth than an “extended” opportunity. Perhaps it has only one or two connected parts. An example of a learning encounter for all teachers in the study that rated as “some opportunity to learn” occurred during the first follow-up session of the school year concerning the practice of *eliciting and using student thinking*. One of the foci for this session was the use of number talks (Parrish, 2016) in the classroom as a way to assist in students' development of fluency with numeric operations at K-5. Within the larger context of number talks, the group discussed how student talk provides the teacher with evidence of thought processes. While engaged in a number talk the teacher elicits and receives evidence of students' ways of thinking about numbers–their flexibility, use of algorithms, etc.,–as well as information about student proficiency with calculations. The teacher uses the evidence gathered to make decisions about questions to ask (or not ask) as well as other appropriate, future number talks. Because this encounter with *eliciting and using student thinking* was of relatively short duration and did not examine multiple elements of using student thinking, the rating was: "some opportunity."

Lastly, in conjunction with either the rating of "some opportunity" or "extended opportunity," notes about the opportunity to learn provided qualitative information concerning the encounter with the practice. The notes recorded specifics from the given learning opportunity. The notes (1) helped in recalling the event and (2) provided evidence for the rating of "some opportunity to learn" or "extended opportunity to learn." Thus, they added to the data collected about the teacher's opportunity to learn.

3.3.2 Instructional quality assessment

Formally titled the Instructional Quality Assessment (IQA) Mathematics Toolkit, "(t)he IQA assesses elements of ambitious instruction in mathematics" (Boston, 2012a, p. 76). The IQA's use for measuring instructional quality in mathematics has been validated in multiple studies over the last decade (e.g., Matsumura et al., 2008; Wilhelm & Kim, 2015). During one of the validation studies, Matsumura and colleagues (2008) found that “as few as two observations might yield a reliable estimate of quality” (p. 292) when using the IQA.

The full IQA Mathematics Toolkit uses both classroom observations and the collection of student assignments to assess instructional quality in mathematics. This study uses only classroom observations. The portion of the IQA dedicated to lesson observations, known as the *IQA Mathematics Lesson Observation Rater Packet, Rubrics, and Lesson Checklist* (Boston, 2012c), contains rubrics for both Academic Rigor and Accountable Talk. This study uses only the Academic Rigor (AR) rubrics. There are five AR rubrics in the IQA. This study employs four of those five rubrics, eliminating the Mathematical Residue rubric (AR-X).

Two trained raters used the four IQA AR rubrics for Potential of the Task, Implementation of the Task, Student Discussion following Task, and Rigor of Teachers’ Questions in this study. The rubric for AR1, Potential of the Task is shown in Figure 3.5, and the set of four AR rubrics used for the study is in Appendix H. The raters used the rubrics at two different junctures in the study: before any coaching activities began and after the conclusion of all coaching activities. Raters evaluated the classroom instruction of each of the ten teachers participating in the study at these two times. Changes in ratings from before coaching to after coaching show changes in the quality of instruction. More specifically, if a teacher chose a task that rated a 2 (low level: procedures without connections) for the Potential of the Task rubric (AR1) at the initial observation, but chooses a task rated as a 4 (high level: doing mathematics) on the same rubric at the final observation, this indicates a shift in ambitious teaching for the effective Mathematics Teaching Practice of *implementing tasks that promote reasoning and problem solving* (NCTM, 2014).

Academic Rigor

RUBRIC 1: Potential of the Task

Did the task have potential to engage students in rigorous thinking about challenging content?

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. • students may be asked to engage in doing mathematics or procedures with connections, but the underlying mathematics in the task is not appropriate for the specific group of students (i.e., too easy <u>or</u> too hard to promote engagement with high-level cognitive demands); • students may need to identify patterns but are not pressed for generalizations or justification; • students may be asked to use multiple strategies or representations but the task does not explicitly prompt students to develop connections between them; • students may be asked to make conjectures but are not asked to provide mathematical evidence or explanations to support conclusions
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 requires no mathematical activity.**
N/A	Students did not engage in a task.

Figure 3.5 Excerpt from the Instructional Quality Assessment Academic Rigor rubric (Boston, 2012c, p. 8)

As another example, consider the rubric for Academic Rigor of the Teacher's Questions, AR-Q. If the initially observed lesson rated a 1, the teacher asked only fact-based or procedural questions that had short or one-word responses (Boston, 2012c). If, however, at the final observation, the teacher gave multiple chances for students to explain their thinking via the questions asked, and the questions helped students connect mathematical ideas (Boston, 2012d), then the final IQA score is a 3 or 4. Scoring a 3 or 4 on AR-Q indicates that the teacher asked "academically relevant questions that provide opportunities for student to elaborate and explain their mathematical work and thinking" (Boston, 2012c, p. 11). The change in rubric score indicates a change in the level of discourse in the mathematics classroom as well as a change in the types of questions asked. This aligns with the effective teaching practice of *posing purposeful questions* (NCTM, 2014). Taken together, higher scores on the AR rubrics provide evidence of more ambitious teacher practices being used in the mathematics classroom.

As described in the paragraphs above, the IQA AR rubrics provide a quantitative score. These scores range from 0, meaning not present, to a high score of 4. Aside from the quantitative scores, the rubrics "represent qualitatively different instructional practices in each of the indicators" (Boston, 2012a, p. 94), so one can use the descriptors at each score level to form a picture of what instructional practice looks like in a given teacher's classroom at the time the observation was done. In other words, the descriptors for each rubric score for a given AR rubric provide qualitative information about the teacher's classroom practices. In sum, Boston wrote that "score levels on the IQA rubrics enable quantitative and qualitative interpretations, as scores represent different levels of instructional quality and specific features of mathematics instruction" (Boston, 2012a, p. 97).

This study uses the eight effective teaching practices for mathematics (NCTM, 2014) as a frame to examine measures of ambitious mathematics teaching. The IQA AR rubrics are well-aligned with a number of the effective Mathematics Teaching Practices. Boston (2012a) has written that the IQA AR rubrics "capture specific aspects of ambitious mathematics instruction:

cognitively challenging tasks, task implementation, students' opportunities to explain their thinking and reasoning during discussions or in written work, and teachers' expectations" (p. 97). Table 3.1 shows the alignment of the AR rubrics with these aspects of the effective teaching practices for mathematics mentioned by Boston. However, the AR rubrics do not explicitly align with all eight of the effective Mathematics Teaching Practices. Documentation of teachers' use of each of the eight effective teaching practices for mathematics is not included the data documented for this study; however, a tool was created to document teachers' use of any of the eight practices during the pre-coaching and post-coaching observations.

Table 3.1 Alignment between the IQA AR rubrics used in this study and the effective Mathematics Teaching Practices

IQA AR rubric	Effective Mathematics Teaching Practice
Potential of the Task (AR1)	Implement tasks that promote reasoning and problem solving
Implementation of the Task (AR2)	Implement tasks that promote reasoning and problem solving
Student Discussion Following Task (AR3)	Facilitate meaningful mathematical discourse
Rigor of Teacher's Questions (AR-Q)	Pose purposeful questions

3.3.3 Effective teaching practices checklist

The Effective Teaching Practices Checklist is aligned to the descriptors for each of the eight effective Mathematics Teaching Practices provided in *Principles to Action: Ensuring Mathematical Success for All* (NCTM, 2014). This checklist potentially provides interesting information about changes in teacher practices aligned to effective mathematics teaching.

However, it does not have the proven reliability and validity of the IQA. An excerpt from the Effective Teaching Practices Checklist is in Figure 3.6. The current version of the full tool is in Appendix I. The Effective Teacher Practices Checklist is designed around the teacher actions associated each effective Mathematics Teaching Practice, therefore, the checklist has eight main sections; one for each of the eight effective mathematics teaching practices. Each effective teaching practice has between four and six teacher actions associated with it in *Principles to Action.* In the first iteration of the tool, the teacher actions were vetted to determine which actions were (1) observable and (2) non-negotiable in ambitious mathematics instruction. With that, the list of teacher actions was narrowed to between two and four for each effective Mathematics Teaching Practice, so each of the eight main sections is divided into two to four subsections based on these non-negotiable descriptors of observable actions. For example, section 4 of the checklist, shown in Figure 3.6 is associated with the effective Mathematics Teaching Practice of *facilitate meaningful mathematical discourse*. The section is divided into two subsections because there are two non-negotiable descriptors of observable teacher actions associated with *facilitating meaningful mathematics discourse*: engaging students in purposeful sharing of mathematical ideas by selecting and sequencing varied students approaches and solution strategies for whole-class analysis and discussion; and ensuring progress towards mathematical goals by making explicit connections to key mathematical ideas in the lesson.

In the next iteration, a rating system was established and descriptors developed for each rating. The checklist assigns a (+), (0), or (-) for each descriptor within each effective Mathematics Teaching Practice based on the observations of teacher actions taken during the lesson. The descriptors within each of the subsections provide concrete guidance to the checklist's user. Information from the checklists can be analyzed according to which teacher actions were exhibited

(+ rating), not exhibited (0 rating), or exhibited in a way that detracted from the intent of the practice (- rating). For example, within the subsection on engaging students in purposeful sharing of mathematical ideas from the section on *facilitating meaningful mathematics discourse*, if the teacher purposefully selects varied student approaches to share during the whole-class debrief, sequences them in a way that can potentially further student progress towards the leaning goal and a mathematically relevant discussion occurs, that would provide evidence of the teacher engaging in the practice of *facilitating meaningful mathematics discourse*. The checklist user records a (+). If on the other hand, any strategies are shared by the teacher, students only share final answers or steps in a procedure, and the teacher questions are in the Initiate-Response-Evaluate (IRE) pattern, then there is evidence the teacher engaged in actions that detract from the practice of *facilitating meaningful mathematics discourse*. The checklist user records a (-). There is a middle ground between these two extremes, wherein the student solutions might be shared but no discussion is engendered, or after students share a solution strategy, the teacher provides this analysis or connection, instead of the students. In that case, the teacher is not fully engaged with the practice, but neither is the teacher detracting from the practice. The checklist user records a (0).

In its current form the Effective Teaching Practice Checklist generates a set of 27 ratings, if used in its entirety. The checklist uses characteristics of the lesson to determine alignment with the non-negotiable descriptors of observable teacher actions that are associated with each of the effective Mathematics Teaching Practices. The Effective Teaching Practices Checklist is meant to provide information about teachers use of or engagement with all eight of the effective teaching practices for mathematics.

For this study, the Effective Teaching Practices Checklist was intended for use at the same times as the IQA AR rubrics were used –before coaching began and after coaching concluded. The checklist was meant to add to the qualitative data provided by the IQA AR rubrics. It was not meant for quantitative data gathering. Although the IQA provides both quantitative and qualitative data regarding the use of ambitious mathematics teaching practices (Boston, 2012a) and is aligned to a subset of the eight effective Mathematics Teaching Practices (see Table 3.1 for this alignment), the Effective Teaching Practices Checklist provides additional information directly aligned to every one of the effective teaching practices in mathematics. Thus, the Effective Teaching Practices Checklist can provide a measure of ambitious mathematics teaching (M. S. Smith, Steele, et al., 2017). This qualitative analysis, when paired with the qualitative components of the IQA AR rubrics might provide insight into patterns in which coached teachers engage that are different from or the same as comparison teachers. For example, evidence of a teacher's use of the effective Mathematics Teaching Practice of *facilitating meaningful mathematics discourse* (NCTM, 2014) includes "Engaging students in purposeful sharing of mathematical ideas by selecting and sequencing varied student approaches and solution strategies for whole-class analysis and discussion" as shown in Figure 3.6. An example of a qualitative pattern that might emerge from the data in the Effective Teaching Practice Checklist is that most or all coached teachers exhibit this action in the Spring, which would rate as a (+), in contrast to the Fall, when some exhibited only the teacher's idea around a single representation, which would rate as (0).

In sum, the Effective Teaching Practices Checklist can provide information about whether, how, and to what extent teachers employ the effective mathematics teaching practices in the observed lesson. However, development of the Effective Teaching Practices Checklist stopped at the second iteration. The tool was not validated within the context of this study. Data from the tool could not be used for analysis of changes in coached teacher actions, nor could the tool be used to qualitatively compare coached teacher actions to comparison teacher actions before coaching began or after coaching concluded. The author hopes to pursue development of this tool in the future to enable a full examination of all eight effective Mathematics Teaching Practices.

Effective Teaching Practice 4: Discourse

Facilitate meaningful mathematical discourse. Effective teaching of mathematics facilitates discourse among students to build shared understanding of mathematical ideas by analyzing and comparing student approaches and arguments. Note: "The discourse in the mathematics classrooms gives students opportunities to share ideas and clarify understanding, construct convincing arguments regarding why and how things work, develop a language for expressing mathematical ideas, and learn to see things from other perspectives" (NCTM, 2014, p. 29).	
Code	What are *teachers* doing?
	• Engaging students in purposeful sharing of mathematical ideas by selecting and sequencing varied student approaches and solution strategies for whole-class analysis and discussion. To earn a +: ○ The teacher purposefully selects varied student approaches to share during the whole-class debrief and sequences them in a way that can potentially further student progress towards the leaning goal. ○ Students share the strategies, and a mathematically relevant discussion occurs. To each a 0: ○ Student solution strategies are shared, and may even be shared in a pre-selected and sequenced way, but no discussion is engendered by the sharing, so no analysis or discussion of the strategies takes place. OR ○ Students do share solution strategies in some random order with only the teacher's voice providing any analysis, instead of the student voices To earn a -: ○ No student solution strategies are shared. Any student sharing is that of answers, next steps in the procedure, or are in the IRE pattern of questioning. OR ○ Student solutions beyond that described in the previous bullet are shared only by the teacher.
	• Ensuring progress toward mathematical goals by making explicit connections to the key mathematical ideas in the lesson. To earn a +: ○ Students are encouraged, via questions or other means, to connect their thinking about the day's task to the mathematical learning goal (or key mathematical ideas, if no learning goal is provided). To each a 0: ○ Only the teacher connects the solution strategies to each other or to the important math of the day. There is no evidence that students have made the connections between the day's work and the important math ideas therein. To earn a -: ○ There are no connections made to the key mathematical ideas, even if discussion does occur.

Adapted from (NCTM, 2014, p. 35)

Coding: + → present in the form described in the bullet point(s)
0 → not fully engaged with the descriptor(s) for a (+) rating in this practice
- →detracts from the practice or from its intent

Figure 3.6 Excerpt from the Effective Teaching Practices Checklist

3.4 Data analysis

3.4.1 Research question 1

To respond to research sub-question 1: *What is the impact on teachers' opportunities to engage with effective teaching practices when content-focused coaching is added to professional development?*, the Opportunities to Learn about Effective Teaching Practices (OtL-ETP) was employed in conjunction with each coaching cycle and each follow-up PD session. The null hypothesis related to this question was, H_0: Coached teachers did not have more opportunities to engage with the effective mathematics teaching practices than comparison teachers. The alternative hypothesis was, H_a: Coaching sessions provide teachers with more opportunities to engage with the effective mathematics teaching practices than participation in outside-the-classroom, school year follow-ups alone. The OtL-ETP tool provides quantitative data in supporting or rejecting the alternative hypothesis. The tool also provides qualitative data to describe the variety of encounters teachers have throughout the study with the effective mathematics teaching practices.

3.4.1.1 Quantitative data

Quantitative data from the OtL-ETP tool generates tables of values summarizing teachers' encounters with the effective Mathematics Teaching Practices. In turn, the data summarized in the tables generates visual displays (e.g. bar graphs) to aid in analysis of similarities and differences between the two groups of teachers and among the group of coached teachers. One table of values generated from the OtL-ETP tool provides data from each follow-up session and each coaching cycle resulting in the total number of opportunities to learn about effective Mathematics Teaching

Practices for each teacher in the study. Note that comparison teachers only had encounters with the effective Mathematics Teaching Practices via the follow-up PD sessions, since the teachers in the comparison group did not engage in coaching. An empty sample of such a table is shown as Table 3.3. This provides a profile for learning opportunities each teacher had and when s/he had the opportunities.

Table 3.2 Data generated from the Opportunities to learn about effective teaching practices tool

	Follow Up PD Session				Coaching Cycle				
	1	2	3	4	1	2	3	4	Total
Comparison Teacher 1									
2									
3									
4									
5									
Coached Teacher 1									
2									
3									
4									
5									

***Note*: Comparison teachers did not take part in coaching. Therefore, data from coaching cycles was not generated and is not applicable.**

Data from Table 3.3 translates to a graph with the horizontal axis having a listing of each teacher from Comparison Teacher 1 through Coached Teacher 5 and vertical axis being a count of effective Mathematics Teaching Practices encountered. A "baseline" bar is displayed across the set of teachers showing the number of effective Mathematics Teaching Practices encountered via the follow-up PD sessions. Each coached teacher's bar then extends above this baseline, showing the number of additional effective mathematics teaching practices s/he encountered via the coaching cycles. A sample of such a bar graph is in Figure 3.7.

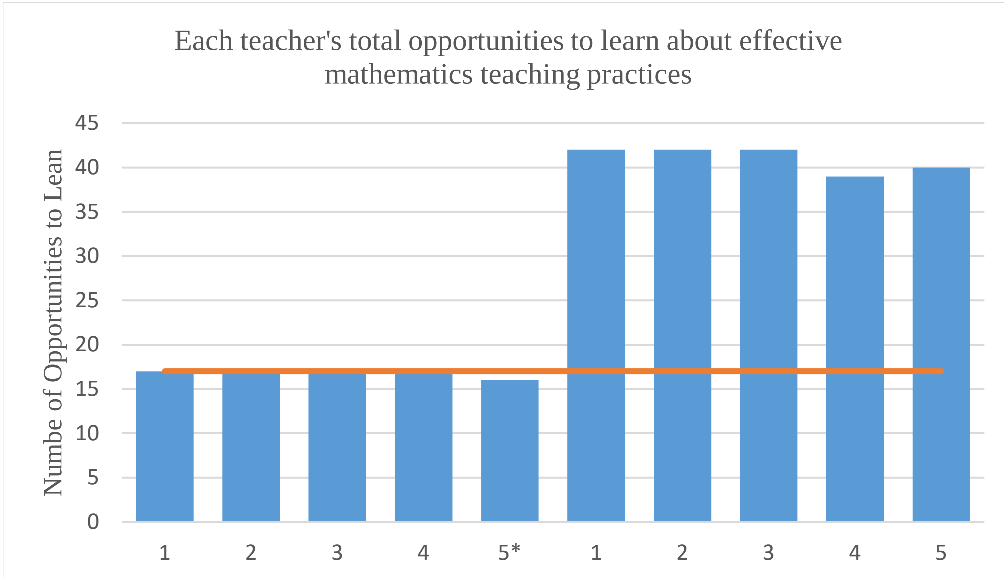


Figure 3.7 Sample graph displaying each teacher's total opportunities to learn about effective Mathematics Teaching Practices

Another table of values generated from the OtL-ETP records the depth of the learning opportunity for the teacher. Because each learning opportunity is not only related to one or more effective mathematics teaching practices, but also rated as “some” or “extended,” the initial data from Table 3.3 can be extended to take the form of Table 3.4. This table separates the total number of opportunities to learn about the effective mathematics teaching practices into those rated as “some opportunity” and those rated as “extended opportunity.” This data produces a stacked or a side-by-side bar graph enabling the comparison of total opportunities to learn as well as comparison of the depth of teachers’ opportunities to learn.

Table 3.3 Data generated from the OtL-ETP tool, and separated by depth of learning opportunity

	Some	Extended	Total
Comparison teacher 1			
2			
3			
4			
5			
Total: Control Teachers			
Coached teacher 1			
2			
3			
4			
5			
Total: Coached Teachers			

Data from Table 3.4 generates additional bar graphs. One such graph displays the sum of the learning experiences rated as "extended opportunity to learn" for control teachers next to that same sum for coached teachers and the sum or "some opportunity to learn" for control teachers next to that sum for coached teachers. This provides a side-by-side comparison of the depth of the learning opportunities for the two sets of teachers in the study. A sample of such a graphical display is shown below as Figure 3.8.

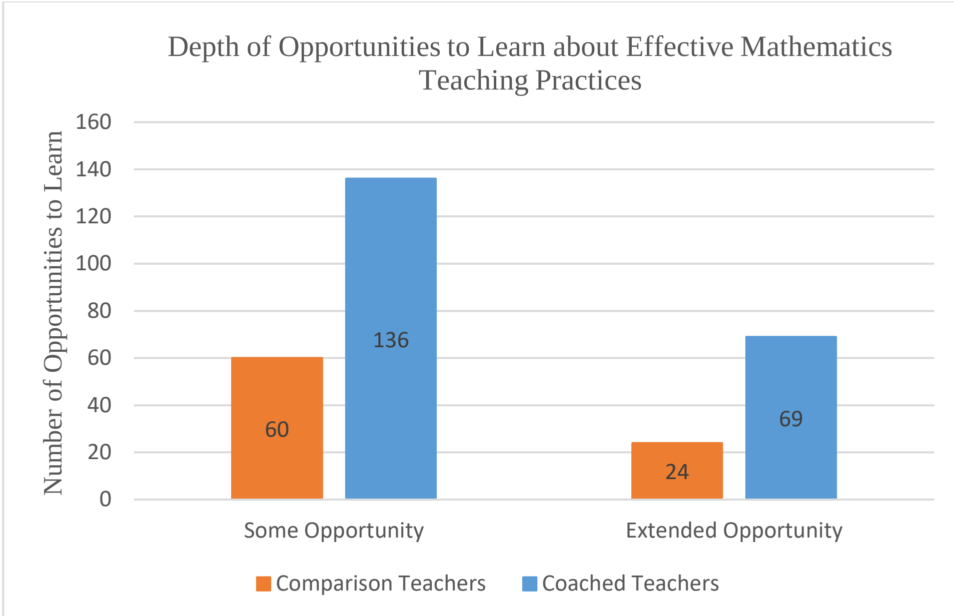


Figure 3.8 Sample graph comparing the depth of control teachers' opportunities to lean about effective mathematics teaching practices with the depth of coached teachers' opportunities to learn

A third table generated from the OtL-ETP data records the opportunities to learn about each of the eight effective Mathematics Teaching Practices examined individually. This table reveals patterns in where each practice was most encountered. These underlying patterns or connections might not be evident from the individual teacher data. Such a table is shown as Table 3.5. Data from this table also generates a number of graphical displays. One such display is a side-by-side bar graph like that described in association with Table 3.4. Instead of the depth of opportunity to learn, each set of two bars refers to one of the eight effective mathematics teaching practices. A sample of such a graphical display is shown in Figure 3.9.

Table 3.4 Opportunities to learn about effective mathematics teaching practices organized by practice

	Effective mathematics teaching practice								Sum of
	Goals	Tasks	Rep'n	Disc	Ques	Conc	Strug	St. Th	ETPs
Comparison teacher 1									
2									
3									
4									
5									
Total: control teachers									
Coached teacher 1									
2									
3									
4									
5									
Total: coached teachers									

***Note*: Column headings are abbreviations for each of the effective mathematics teaching practices chose by the author. Goals => Establish mathematics goals to focus learning; Tasks=> Implement tasks that promote reasoning and problem solving; Rep'n => Use and connect mathematics representations; Disc => Facilitate meaningful mathematics discourse; Ques => Pose purposeful questions; Conc => Build procedural fluency from conceptual understanding; Strug => Support productive struggle in learning mathematics; St. Th => Elicit and use evidence of student thinking.**

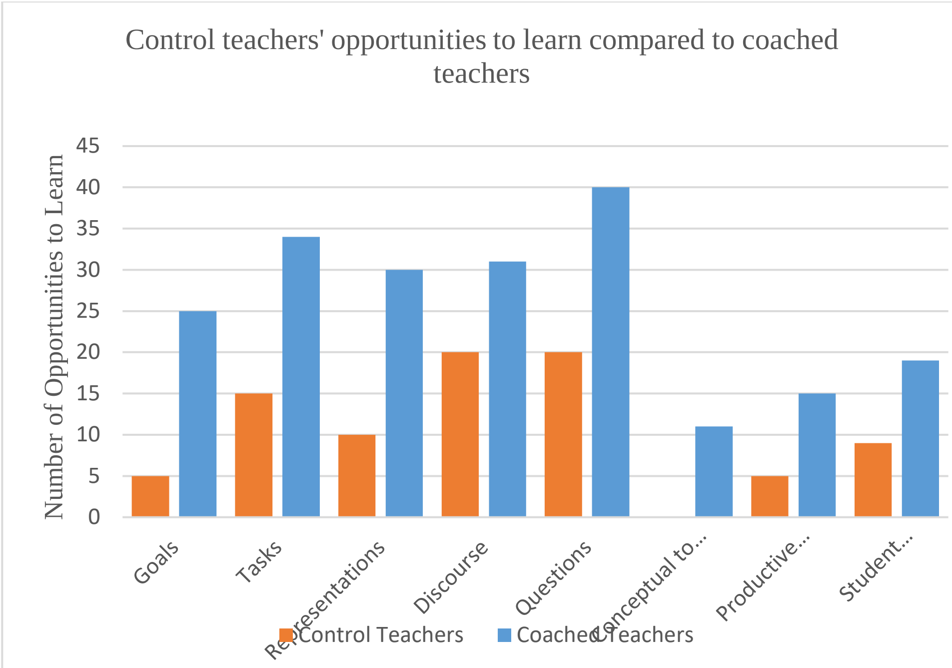


Figure 3.9 Sample graphical display contrasting control teachers' opportunities to learn about each of the eight effective mathematics teaching practices with coached teachers' opportunities

These tables show whether coaching focused on different effective Mathematics Teaching Practices than the follow-up PD sessions and whether coaching increased the degree or depth ("some" vs. "extended") of opportunities to learn. From the three tables and graphs generated via the OtL-ETP data, it became obvious that coached teachers had more opportunities for learning about the effective Mathematics Teaching Practices than did comparison teachers. It was anticipated that coached teachers would have a higher overall total of opportunities to learn, more numerous opportunities to engage in *each* of the eight practices, as well as a higher total of "extended" opportunities to learn.

3.4.1.2 Qualitative data

The last column of the OtL-ETP data collection tool contains notes about a given coaching cycle or follow-up PD session. The notes provide evidence or proof that the teacher did have an opportunity to learn about a particular effective mathematics teaching practice during the coaching cycle or PD session. For example, evidence showing that the teacher had the opportunity to learn about *establishing mathematics goals to focus learning* was "Coach created three possible goals statements related to the topic for the lesson. Together the coach and teacher analyzed each and chose the learning goal for the student lesson."

The notes provide a qualitative component to the OtL-ETP data collection tool. They provide information about what a given learning opportunity looked like, from the coach's perspective. This qualitative component helps differentiate between an "extended" opportunity to learn and a learning opportunity rated as "some." Thus, the notes help clarify the difference between the two possible ratings, potentially providing a "tipping point" of sorts between a learning opportunity that has less depth and one that has more depth.

Additionally, the qualitative component of the notes helps clarify which of the eight effective Mathematics Teaching Practices teachers encountered. For example, an opportunity to learn about a particular teaching move like asking an assessing question followed by a student response and then an advancing question (M. S. Smith & Stein, 2011) is attributed to an opportunity to learn about the practice of *posing purposeful questions*, *supporting productive struggle*, *eliciting and using student thinking*, or some combination thereof. The notes taken in conjunction with this learning opportunity provide justification for the choice of teaching practice(s) encountered. Without the notes, the coach might default to recording this opportunity to learn as dealing only with posing purposeful questions. While this might be the case, it might also be the case that the learning opportunity revolved more around eliciting and using student thinking than it did around questioning alone.

3.4.2 Research question 2

The Instructional Quality Assessment (IQA) Academic Rigor (AR) rubrics respond to the second research question: *What is the impact on teachers' **use** of ambitious teaching practices when content-focused coaching is added to professional development?*. There were two null hypothesis associated with this research question: H_{01}: Coached teachers' use of ambitious teaching is no different than comparison teachers', and H_{02}: Coached teachers' use of ambitious teaching remains the same from the study's beginning to the study's end. The two alternative hypotheses associated with the question were: H_{a1}: Coached teachers employ more ambitious teaching than comparison teachers, and H_{a2}: Coached teachers increase their use of ambitious teaching from the beginning of the study to the end.

3.4.2.1 Quantitative data

The author and an outside evaluator used four of the IQA AR rubrics to assess the ten teachers' use of ambitious teaching practices. The author and outside evaluator used the rubrics before the start of any coaching activities and after the conclusion of all coaching activities. The IQA AR rubric scores were used (a) to compare the coached teacher group to the comparison teacher group and (b) to examine any changes in the scores for the coached teacher group from before coaching begins to after coaching concludes. The four rubrics from the IQA used in this study were AR1: Potential of the Task, AR2: Implementation of the Task, AR3: Student Discussion following Task, and AR-Q: Rigor of Teachers' Questions. Appendix H shows the four rubrics. Before the start of any coaching activities, the two evaluators performed consensus scoring for each of the teachers in the study on one lesson. After the conclusion of all coaching activities, the two evaluators again performed consensus scoring for each of the ten teachers. Data from the evaluators were summarized in data tables akin to Table 3.6.

Table 3.5 IQA consensus scores

		Fall 2017					Spring 2018				
		AR1	AR2	AR3	AR-Q	Total	AR1	AR2	AR3	AR-Q	Total
Comparison teachers	1										
	2										
	3										
	4										
	5										
Coached teachers	1										
	2										
	3										
	4										
	5										

Data from the IQA scores were also be shown with a bar graph to compare cumulative starting to ending scores or with stacked bar graph to compare individual AR rubric scores as well as cumulative scores before and after coaching. It was anticipated that coached teachers' cumulative scores on the IQA would increase from Fall 2017 to Spring 2018 as would their scores on each of the AR rubrics. To determine if the composite scores on the IQA changed significantly from Fall to Spring for the coached teachers, the Wilcoxan Signed-Rank Test (nonparametric paired t-test) was used. The Wilcoxan Signed-Rank Test was also used to determine if scores on the AR rubrics change significantly. As the name says, this test uses signed ranks instead of absolute scores. To determine significance of results, the signed ranks are summed, so the rank order of teachers' scores, along with whether scores increased or decreased from Fall 2017 to Spring 2018, were important for this test.

Whereas there were a total of 20 AR rubric scores to compare from the beginning of the study to the end, there were only five composite IQA scores to compare. With the lower number of scores for the composite IQA, there were only five signed ranks to add. With only five composite scores, *all* the composite IQA scores needed to either stay the same or increase for the sum of the signed ranks to be significant. With 20 AR rubric scores to compare, a limited number of the signed ranks can be negative, indicating a decrease in AR rubric score from fall to spring, and the Wilcoxan Signed-Rank test still provides a significant result. A significant result for both of these tests indicates that coached teachers changed the nature of their instruction away from traditional pedagogies and towards more ambitious teaching practices. If the only significant results comes from comparing the AR rubric scores, then the teachers in the coached group either increased their scores on certain rubrics or certain teachers increased scores while others did not. This means the group of teachers only changed some of the teaching practices measured by the IQA AR rubrics. Recall from Table 3.6 that those are: *implement tasks that promote reasoning and problem solving*, *facilitate meaningful mathematical discourse*, and *pose purposeful questions*. While it is certainly a better outcome for both results to be significant, having only AR rubric scores increase significantly still indicates a shift in practice. Boston and Smith (2009, 2011) saw significant changes in teacher practice as a result of the ESP professional development while examining only task selection and implementation. This study examined task selection and implementation as well as discourse and teachers' use of questions. Seeing a significant increase on composite IQA scores and an insignificant change on AR rubric scores was possible but unlikely. For this to occur, all coached teachers' composite IQA scores needed to increase or stay the same while enough AR rubric scores decreased to make that change insignificant. If this would have occurred, a closer examination of changes in AR rubric scores would have been warranted. If the decreased scores were still considered high scores (meaning decreases from 4 to 3), then the indication from overall composite IQA scores of increased ambitious practice would be heeded. However, if AR rubric scores decreased rom from high scores to low scores (3 to 2) or decreased from low to lower scores (2 to 1), then the AR rubric outcome would take precedence, meaning that the group of coached teachers had not adopted more ambitious practices.

It was also anticipated that while coached and comparison teachers' scores were comparable in Fall 2017, they would be different in Spring 2018, with coached teachers outscoring comparison teachers. All composite scores from both coached and comparison teachers were compared in the Fall using the Mann-Whitney test (nonparametric unpaired t-test) to determine if there were differences between the two groups at the onset of the experiment. All composite scores for coached and comparison teachers were again compared in the Spring to determine if there were differences then. Like the Wilcoxan Signed-Rank test, the Mann-Whitney test uses rank order. Unlike the Wilcoxan Signed-Rank test, the Mann-Whitney test does not attach signs to the rankings. For one group's scores to be statistically different from the other's, the sum of their ranks (1 through 10 in this case) has to be significantly less or significantly more than the sum of the ranks for the other group. If results show what was expected, then the sum of the coached teachers' rankings at the beginning of the study are close to the sum of the comparison teachers' rankings, but the sums of rankings would diverge at the conclusion of the study, with coached teachers rankings being better. This would have shown that the coached teachers and comparison teachers started with similarly ambitious or traditional methodologies but the teacher group with the significantly higher sum of ranks (presumably, the coached teachers) ended with more ambitious practice exhibited. However, if the sum of ranks was close at the beginning, and the sum of ranks was still close at the conclusion of the study, that indicates that the two groups of teachers had similarly ambitious practices at the beginning of the study and had similarly ambitious practices at the conclusion of the study. In this case, the scores warrant a close examination. If closer examination shows that *both* groups increased their scores, maintained their scores, or decreased their scores from before to after coaching, then a better interpretation of results may be obtained from comparing the coached group to itself with the aforementioned Wilcoxan Signed-Rank test.

Data from Table 3.6 was also used to track the number of high-level ratings–ratings of 3 or 4 according to Boston (2012a)–each teacher received in the Fall and then in the Spring. The number of high-level ratings from coached and comparison teachers from Fall and from Spring was then compared via a chi-squared test. The chi-squared test tells if there is an association between being coached and the number of high-level ratings at two differing points in time. Table 3.7 shows a table organized for showing the data used in the chi-squared test. If the number of high-level rankings was greater than "expected" for the coached teachers, in other words, if the result of the chi-squared test is significant, that means that coached teachers' practices, as measured on the IQA, were more ambitious than the comparison teachers' practices in the Spring 2018.

Lastly, composite scores from the four AR rubrics for classroom observation were grouped according to ranges to facilitate the use of a variation of the chi-squared test for data sets where the expected valued may be less than or equal to five. This test is called the Fisher Exact Probability test. The test tells if there is an association between being a coached teacher and getting lower or higher composite IQA scores. Lower composite scores are those less than nine. Higher composite scores are greater than or equal to nine up to the highest total possible score of 16. A breaking point of nine is used because teachers who earned a composite score of at least nine had to receive a high-level rating of 3 or 4 on at least one of the four AR rubrics in use. The Fisher Exact Probability test was run for the Fall 2017 data and again for the Spring 2018 data. A table like 3.8 was used to summarize each data set according to ranges of scores. If the number of high-level composite scores was greater than "expected" for the coached teachers, in other words, if the result of the Fisher Exact Probability test is significant, then like the results of the chi-squared test just described, that means that coached teachers' practices, as measured by the composite IQA scores, were more ambitious than the comparison teachers' practices.

Table 3.6 Summary table showing number of high-level ratings for each use of the IQA AR rubrics

	Fall 2017	Spring 2018	Total
Coached teachers			
Comparison teachers			
Total			

Table 3.7 Summary of composite scores for the four AR rubrics from the IQA used for this study

	$0 \le$ composite < 9	$9 \le$ composite ≤ 16	total
Coached Teachers			
Comparison Teachers			
Total			

3.4.2.2 Qualitative data

While the IQA provides the quantitative data for responding to the second research question in this study, the IQA is also a qualitative tool. As Boston, Bostic, Lesseig, and Sherman (2015) share, "The IQA score levels are also very descriptive, indicating specific characteristics or frequencies of instructional practice necessary for each score level. The detailed descriptors for each score level and the consistency in score levels across rubrics facilitate qualitative interpretations of the IQA results" (p. 159). The qualitative analysis derived from the IQA AR rubrics provides insight into patterns in which coached teachers engage that are different from or the same as comparison teachers. For example, the descriptor for AR-Q: Rigor of Teachers' Questions, level 4 states that "the teacher consistently asks academically relevant questions that provide opportunities for students to elaborate and explain their mathematical work and thinking,…identify and describe important mathematical ideas in the lesson, or make connections between ideas, representations, or strategies" (Boston, 2012b, p. 8). The level 2 descriptor for this same rubric states that "There are one or more superficial, trivial, or formulaic efforts to ask academically relevant questions probing, generating discussion, or exploring mathematical meanings and relationships" (p. 8). Beyond a rating being high or low on its face, these descriptors that are paired with the numeric value allow the coach or researcher insight into what is actually happening in the classroom regarding a teacher's questioning patterns per se. The researcher can qualitatively compare the descriptions aligned with the coached teachers' scores with the descriptions for the rubric scores earned by the comparison teachers. This comparison yields insights regarding the asking of which academically relevant questions. Asking questions that allow students to elaborate and explain their mathematical thinking aligns with the effective Mathematics Teaching Practice of *posing purposeful questions* (NCTM, 2014). While this qualitative information is gleaned from the same tool as the numeric scores, this information is different from a numeric score because it provides a window into the classroom regarding effective mathematics instruction.

Similar insights are gleaned from a comparison of coached teachers' score descriptors from the fall and coached teachers' score descriptors from the spring. The qualitative analysis provides information about changes in coached teachers practices that align with IQA AR rubric descriptors as well as the descriptors for the related effective Mathematics Teaching Practice. For example, *Principles to Action* (NCTM, 2014) describes the practice of *facilitating meaningful mathematical discourse* by stating, "Effective teaching of mathematics facilitates discourse among students to build shared understanding of mathematical ideas by analyzing and comparing student approaches and arguments" (p. 10). *Facilitating meaningful mathematical discourse aligns* with AR3: Student Discussion Following Task. The descriptors for the low level ratings of 1 or 2 on the AR3 rubric include information about students either showing procedural work or providing only brief responses to teacher questions. The descriptors for the high level ratings of 3 or 4 on this rubric include information about the quality of students' sharing of their mathematical work and engagement in a discussion that is student-led or teacher guided. The high level descriptors sound similar to the description of *facilitating meaningful mathematical discourse.* So, if coached teachers earned more high level ratings for AR3 in the spring than they did in the fall, the qualitative nature of the IQA provides evidence that the classroom teacher's practice has become more ambitious. Again, this goes beyond the numeric score and provides the coach with an awareness of ambitious practice. Taken together, the IQA AR rubrics provide information about whether, how, and to what extent the observed teacher employed ambitious teaching practices aligned with *implementing tasks that promote reasoning and problem solving*, *facilitating meaningful mathematical discourse*, and *posing purposeful questions* in the observed lesson.

3.5 Summary

This chapter outlined the approach for the proposed study. The chapter described the design of the study including information about the participants and the coaching model. Special attention was given to explaining the phases of the coaching cycle. Next, the chapter described the junctures at which data was collected: during follow-up PD, via the classroom observations before the start and after the conclusion of the coaching cycles, and via the coaching cycles themselves. Then, the chapter discussed the two main tools used in the study and one tool created for the study but not validated for use. Those tools employed in the study were the Opportunities to Learn about Effective Teaching Practices data collection tool (OtL-ETP) and the Instructional Quality Assessment (IQA). The unvalidated tool was the Effective Teaching Practices Checklist. Finally, the last section of the chapter summarized the data analysis procedures for answering each research question, including both qualitative and quantitative methods

4.0 Results

This chapter presents the research results of the study described in the previous chapter. The chapter attempts to answer the questions:

- How does proximal, in situ professional development in the form of content-focused coaching paired with outside-the-classroom professional development facilitate a change in mathematics teachers' pedagogical practices from traditional to more ambitious in nature?
    a. What is the impact on teachers' *opportunities to learn about* ambitious teaching practices when content-focused coaching is added to professional development?
    b. What is the impact on teachers' *use of* ambitious teaching practices when content-focused coaching is added to professional development?

The results are organized by research question. Within the results for the first research question, overall opportunities to learn are analyzed first, followed by the depth of opportunity to learn and finally an analysis pertaining to each of the eight effective Mathematics Teaching Practices is provided. Quantitative results precede qualitative results in each of these subsections. Quantitative results also precede qualitative results in responding to the second research question.

4.1 Opportunities to learn about ambitious teaching

The Opportunities to Learn about Effective Teaching Practices data collection tool (OtL-ETP) (see also Appendix G), as described in chapter 3, tracked the occasions when teachers involved in the study had the chance to learn about one or more of the practices associated with ambitious teaching. Recall that this study uses the effective Mathematics Teaching Practices (National Council of Teachers of Mathematics [NCTM], 2014) to exemplify practices associated with ambitious teaching in mathematics. The OtL-ETP recorded teachers' opportunities to learn about the effective Mathematics Teaching Practices during (1) each of the four coaching cycles for the five coached teachers and (2) each of the four follow-up professional development (PD) sessions spread throughout the 2017-18 school year[8]. Data gathered for the follow-up PD on K-5 mathematics teaching applies to the five coached teachers and the five comparison teachers. The *overall* data is analyzed in section 4.1.1. Following that, in section 4.1.2, the *depth* of teachers' opportunities to learn is analyzed. Lastly, in section 4.1.3, data analysis pertaining to *each* of the eight effective Mathematics Teaching Practices is provided. Within each section, qualitative data is examined after quantitative data.

4.1.1 Overall opportunities to learn about effective Mathematics Teaching Practices

4.1.1.1 Quantitative results

Because data from the follow-up PD sessions applies to all ten teachers in the study, there exists a set of common opportunities to learn about the effective Mathematics Teaching Practices. In three of the four follow-up PD sessions, teachers had opportunities to learn about four distinct effective Mathematics Teaching Practices, while the last of the follow-up PD sessions had five

[8] As mentioned in chapter 3 of this document, all teachers participating in the outside-the-classroom PD attended ten days of content-focused PD during Summer 2017. While those ten days focused on content knowledge and specialized content knowledge, there was the implicit involvement of pedagogical practices during the summer.

such opportunities. Thus, all teachers[9] had 17 opportunities to learn about the effective Mathematics Teaching Practices during the follow-up PD sessions for the 2017-18 academic year. That is the sum total of opportunities to learn about the effective Mathematics Teaching Practices for the comparison teachers. Table 4.1 contains this data in the columns headed by "Follow-up PD Session." See Appendix J for a listing of which effective Mathematics Teaching Practices were encountered in each of the four follow-up PD sessions.

Table 4.1 Number of opportunities to learn about effective Mathematics Teaching Practices by session

		Follow-Up PD Session				Coaching Session				
		1	2	3	4	1	2	3	4	Total
Comparison teachers	1	4	4	4	5	n/a	n/a	n/a	n/a	17
	2	4	4	4	5	n/a	n/a	n/a	n/a	17
	3	4	4	4	5	n/a	n/a	n/a	n/a	17
	4	4	4	4	5	n/a	n/a	n/a	n/a	17
	5	4	4	4	4	n/a	n/a	n/a	n/a	16
Coached teachers	1	4	4	4	5	4	6	7	8	42
	2	4	4	4	5	5	7	7	6	42
	3	4	4	4	5	7	6	7	5	42
	4	4	4	4	5	4	5	5	8	39
	5	4	4	4	5	4	5	8	6	40

***Note:* Comparison teacher 5 was absent from an afternoon of follow-up PD and missed an opportunity to learn about one of the effective Mathematics Teaching Practices. Therefore, this teacher's total number of opportunities to learn is one less than that of the other four comparison teachers.**

[9] Control teacher 5 was absent from an afternoon of follow-up PD and missed an opportunity to learn about one of the effective mathematics teaching practices. Therefore, this teacher's total number of opportunities to learn is one less than that of the other four control teachers. This is also reflected in tables and figures to follow.

Figure 4.1 shows the 17 common opportunities to learn as a horizontal segment atop the bars for comparison teachers. Table 4-1 and Figure 4-1 show coached teachers had more exposures to the effective Mathematics Teaching Practices compared to comparison teachers.

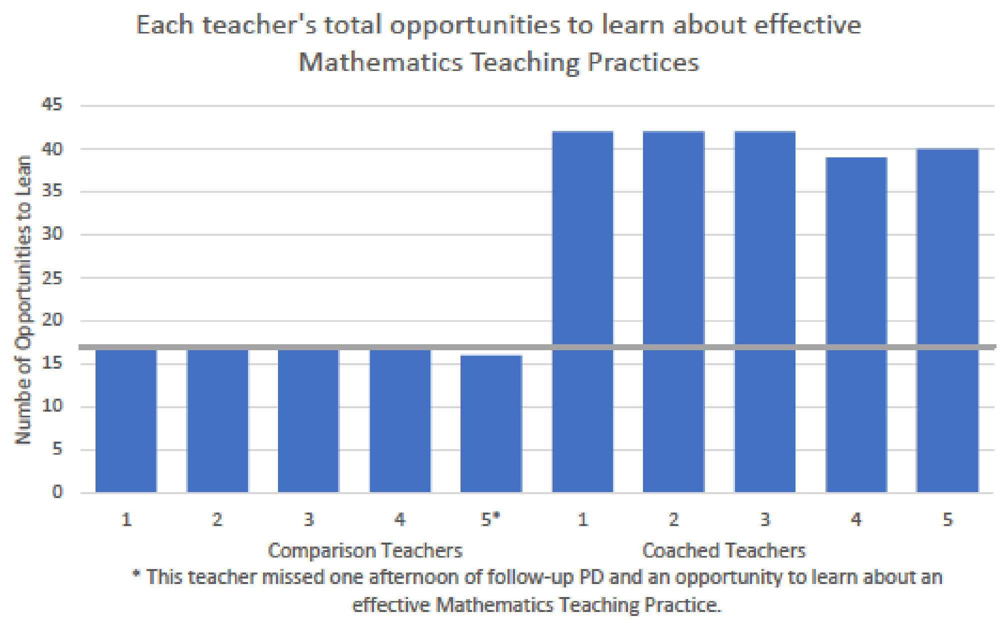


Figure 4.1 Opportunities to learn about effective Mathematics Teaching Practices by teacher

Cumulatively, the group of five comparison teachers had 84 opportunities to learn about effective Mathematics Teaching Practices, while the group of five coached teachers had 205 such opportunities. Each comparison teacher had 17 opportunities to learn about the practices, except for one teacher, who had 16 chances. In comparison, coached teachers had an average of 41 chances to learn about the same practices during the 2017-18 school year, with a range of 39 to 42 opportunities per teacher.

Each teacher who received coaching had more than two times the number of chances to learn about the practices critical to ambitious mathematics teaching compared to any teacher in the comparison group. This finding seems logical based on the number of interactions each teacher had with coaching and professional development within the school year. Coached teachers interacted directly with the coach or facilitator for professional development eight times: four during coaching cycles and four during the follow-up PD. During the same period, comparison teachers interacted with the facilitator four times; only during the follow-up PD episodes.

4.1.1.2 Qualitative results

All participants in the study had access to quality professional development both during the ten summer days and during the four follow-up sessions because the common PD experiences were focused on content knowledge, involved active learning, and were coherent with mathematics standards (Garet et al., 2001). As mentioned above, some of these common PD experiences from the follow-up sessions were directly connected to the effective Mathematics Teaching Practices. One such experience connected to the practice of *facilitating meaningful mathematical discourse*. At this follow-up PD session, participants first read the *Case of Ms. Bouchard* (Huinker & Schefelker, 2016). Next, participants received information about the effective Mathematics Teaching Practices from *Principles to Action* (NCTM, 2014) and discussed where Ms. Bouchard had facilitated discourse. This led to discussion of the *5 Practices for Orchestrating Productive Mathematics Discussions* (M. S. Smith & Stein, 2011). Then, participants engaged in using the *5 Practices* with a task called *Maria's Money* (The Charles A. Dana Center at the University of Texas at Austin, n.d.) (See Appendix K.1). For the task, they anticipated student solution strategies, devised assessing and advancing questions to be used during monitoring, and used sample student work to decide which questions to ask.

While the previous example applies to all teachers involved in the study, the data shows that coached teachers had many additional opportunities to learn about the effective Mathematics Teaching Practices. Some of the added opportunities had commonalities across the group of coached teachers. For example, the first coaching session for each teacher contained an opportunity to learn about *establishing mathematics goals to focus learning*. In advance of the planning session, the coach created three possible goal statements related to the topic for the lesson. The coach shared the possible goals with the teacher, and together the coach and teacher analyzed the strengths, weaknesses, and affordances of each. Together, they chose the goal that provided the best alignment to the desired learning. The coach provided some information gleaned from *Taking Action: Implementing Effective Mathematics Teacher Practices in Kindergarten-Grade 5* (Huinker & Bill, 2017) to compare and contrast learning goals to performance goals. Then, the coach and teacher turned to their monitoring charts to refine the goal statements each had initially provided before meeting face-to-face. By the conclusion of the planning discussion, the teacher-coach pair collaboratively crafted a learning goal appropriate for the upcoming lesson. Following the implementation of the lesson, when the teacher and coach reconvened to debrief, they discussed whether and how the learning goal informed decisions during the lesson as well as decisions about instructional next steps. This general encounter repeated itself for each of the five coached teachers. The coached teachers' opportunities to learn about goals were qualitatively different from the one opportunity that all teachers had to learn about goals in the context of a follow-up PD session. The goals the coached teachers wrote were immediately relevant to *their* lessons and *their* students' learning. The goals the coached teachers wrote were the actual goals to be attained in short order by the children in their classrooms. In contrast, the encounter with goals from the follow-up PD session dealt with learning goals in connection to a discussion of strategies for effective formative assessment (Wiliam, 2013) wherein "clarifying, sharing, and understanding goals for learning and criteria for success" (p. 1) is the first of five such strategies. The discussion during the follow-up PD session was not immediately relevant a given lesson as was the case in coaching, nor did the teachers write goals for their students during the follow-up session.

The coach purposefully pre-planned for each coached teacher to have a similar encounter with *establishing mathematics goals to focus learning* during their first coaching cycle. However, regardless of the effective Mathematics Teaching Practice(s), most of the additional learning encounters provided during the coaching sessions took on a more unique flavor than the initial opportunity each coached teacher had to learn about *establishing mathematics goals to focus learning*. For example, a unique opportunity to learn about *building procedural fluency from conceptual understanding* presented itself with Coached teacher 3. This teacher implemented a sequence of related tasks. A task in the middle of this series was purposefully scheduled during the last coaching cycle. Because the teacher chose to implement a sequence of tasks, the coach and teacher were able to talk in detail, during the planning session, about whether and how these tasks connected and how the teacher might leverage the tasks to help students work from conceptual understanding towards procedural fluency. Then, in the debrief session, the coach-teacher team discussed that students had not yet made as much progress towards the ultimate learning goal as hoped. This led to a decision to integrate an additional, jointly agreed-upon task in the progression. Finally, the coach and teacher used the remainder of their discussion time to revisit the group's, as well as students' individual trajectories from conceptual understanding towards procedural fluency.

4.1.2 Depth of opportunities to learn about effective Mathematics Teaching Practices

4.1.2.1 Quantitative results

Within the OtL-ETP data collection tool, each opportunity to learn was judged as an "extended" opportunity or as "some" opportunity to learn. An "extended opportunity to learn" is a deep or lengthy encounter with a practice. It may entail multiple connected parts or one in-depth activity or conversation. A learning opportunity judged as "some opportunity to learn" entails something shorter in duration or an encounter with less depth than an "extended" opportunity. Perhaps it has only one or two connected parts. As previously discussed, every teacher attended the four follow-up PD sessions and had 17 common opportunities to learn about the effective teaching practices. Of these 17 common experiences for each teacher, 12 were rated as "some," with the remaining five experiences rated as "extended." Figure 4.2 shows the breakdown of the ratings for the common experiences by follow-up session.

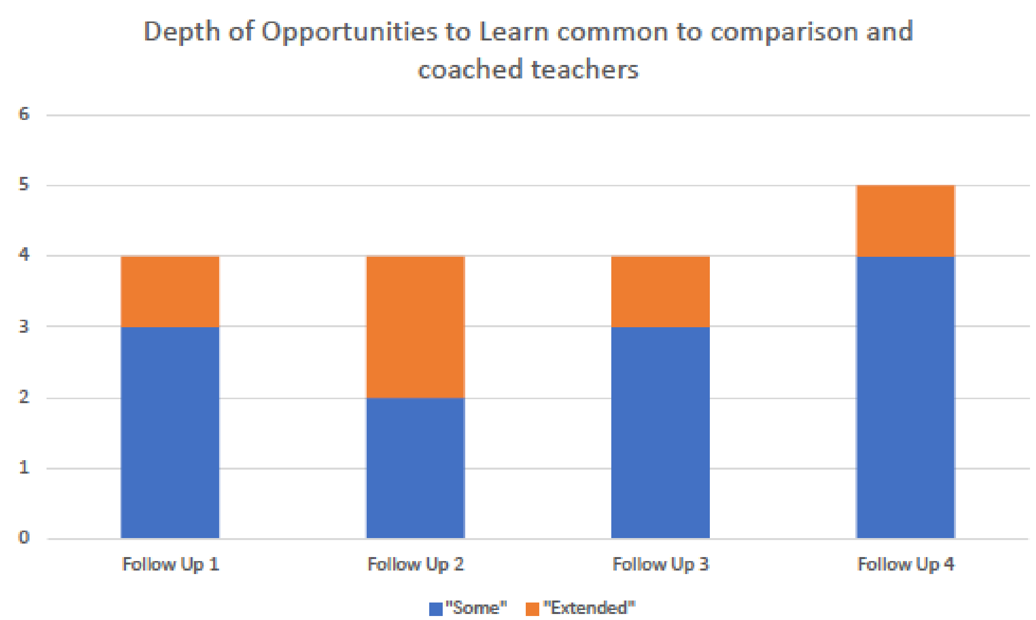


Figure 4.2. Depth of opportunities to learn about effective Mathematics Teaching Practices common to all teachers

Table 4.2 shows the numeric data concerning depth of the learning experiences for each teacher. Cumulative data shows that of the 84 opportunities for the group of comparison teachers to learn about the effective Mathematics Teaching Practices, 60 of them were rated as “extended,” while the remaining 24 were rated as “some” opportunity to learn. Of the 205 opportunities for the group of coached teachers to learn about the effective Mathematics Teaching Practices, 136 were “extended” and 69 were rated as “some.”

Table 4.2 Number of opportunities to learn about effective Mathematics Teaching Practices by depth of learning experience

	Some	Extended	Total
Comparison teacher 1	12	5	17
Comparison teacher 2	12	5	17
Comparison teacher 3	12	5	17
Comparison teacher 4	12	5	17
Comparison teacher 5	12	4	16
Total: Comparison Teachers	**60**	**24**	**84**
Coached teacher 1	28	14	42
Coached teacher 2	29	13	42
Coached teacher 3	29	13	42
Coached teacher 4	25	14	39
Coached teacher 5	25	15	40
Total: Coached Teachers	**136**	**69**	**205**

4.1.2.2 Qualitative results

There were certainly quantitative differences in the overall number of encounters comparison teachers had with the effective teaching practices for mathematics compared to the number of encounters coached teachers had with the same practices. There were also differences in the number of "extended" encounters each group of teachers had with the effective teaching practices for mathematics as well as differences in the encounters rated as "some" opportunity to learn. In addition to the differences in the raw numbers of encounters, there was a qualitative difference in the opportunities provided through coaching.

In planning for the implementation of a task called *Scaling Up and Down* (Illustrative Mathematics, 2016b) (See Appendix K.2), a coached teacher had an "extended" opportunities to learn about the effective Mathematics Teaching Practice of *posing purposeful questions*. In their planning session, the teacher and coach each created and then shared and discussed assessing and advancing questions for each of five different anticipated student solution strategies. Unlike the purposeful questions created by teachers in follow-up PD, the questions created for use during the implementation of this task were employed in this teacher's actual class, with the teacher's actual students. Some questions were used as they were written in the Monitoring Chart, some were modified when used, and some questions were not used during the lesson. After the lesson implementation, the teacher and coach debriefed about the questions' effectiveness in moving students towards the learning goal of understanding the effect of the scalar in a multiplicative expression. Even when comparison teachers had the occasion to create assessing and advancing questions, they never had an opportunity such as the one afforded to this coached teacher, wherein questions were created with the direct support of a more knowledgeable other, implemented with the teachers' own students, and reflected upon after the lesson with the continued support of the coach[10].

---

[10] Throughout the results section, the author selected examples of both coached and comparison teachers' encounters with learning about the effective teaching practices for mathematics. Please note that illustrations for each of the eight effective Mathematics Teaching Practices will not appear in all sections of the results chapter. However,

Even learning opportunities rated as “some” often took on a different flavor during coaching than they did during the follow-up PD. Another illustration provides a sense of the qualitative difference between the encounter for a comparison versus a coached teacher, even when both encounter the same effective Mathematics Teaching Practice rated at the same depth of opportunity to learn. While the effective mathematics teaching practice of *facilitating meaningful mathematical discourse* had been discussed in follow-up PD, this practice had not migrated into all of the coached teachers’ classrooms. One coached teacher in particular was hesitant to use pairs or groups in her class of primary students because she believed the children did not know how to talk to one another. During the debrief portion of the first coaching cycle, the advantages of using student groups to engender mathematical discourse were discussed. Subsequently, in the planning portion of the second coaching cycle, the coach and teacher designed an introduction to discourse for the students. At the beginning of class, the coach and teacher role-played a turn-and-talk for the students who sat on the carpet, attending to the exchange. The coach and teacher then asked the students what they noticed about the exchange and made note of student responses. In the coaching cycles that followed, these students consistently worked in pairs or small groups. Though this opportunity to learn was rated as “some,” it was qualitatively different from an opportunity with the same rating provided to all teachers in the follow-up PD. A learning opportunity afforded to all teacher entailed a discussion of facilitating classroom discourse as it related to the article, “Five ‘Key Strategies’ for Formative Assessment,” (Wiliam, 2007). One of these key strategies discussed in the follow-up PD session was called “Engineering effective classroom discussion, questions, activities, and tasks that elicit evidence of students’ learning” (p. 2). Within this key

---

between chapter 3 and chapter 4 of this document, the reader will find at least one illustration for each of the eight practices.

strategy, teachers are supposed to use or craft formative assessment items and use the student group's responses to engender a discussion during which student views can be aired with justifications. Teachers discussed the example formative assessment item provided in the article along with possible responses students might provide for their justifications. This encounter with *facilitating meaningful mathematical discourse* was different than the one provided for the coached teacher.

4.1.3 Opportunities to learn analyzed by effective Mathematics Teaching Practice

4.1.3.1 Quantitative results

Not only does the OtL-ETP data collection tool code opportunities to learn as "some" or "extended," the tool also codes the data according to which of the effective Mathematics Teaching Practice(s) were encountered within the learning opportunity. What follows is the quantitative analysis of that data. Table 4.3 provides this data, showing the spread of the 17 common opportunities to learn about the effective Mathematics Teaching Practices for all teachers and the differential spreads for each coached teacher. During the follow-up sessions, teachers had the greatest number of opportunities to learn about the practices of *facilitate meaningful mathematics discourse* and *pose purposeful questions*. Teachers had four chances to learn about each of these two practices; one opportunity in each of the four follow-up sessions for a total of 20 opportunities to learn about each of these to effective Mathematics Teaching Practices.

Table 4.3 Number of opportunities to learn about each effective Mathematics Teaching Practice ordered from least to greatest according to the comparison group

	Effective Mathematics Teaching Practice								Sum of ETPs
	Conc	Strug	Goals	St. Th	Rep'n	Tasks	Disc	Ques	
Comparison teacher 1	0	1	1	2	2	3	4	4	17
Comparison teacher 2	0	1	1	2	2	3	4	4	17
Comparison teacher 3	0	1	1	2	2	3	4	4	17
Comparison teacher 4	0	1	1	2	2	3	4	4	17
Comparison teacher 5	0	1	1	1	2	3	4	4	16
Total: Comparison group	**0**	**5**	**5**	**9**	**10**	**15**	**20**	**20**	**84**
Coached teacher 1	2	3	5	5	6	7	6	8	42
Coached teacher 2	2	3	5	4	6	7	7	8	42
Coached teacher 3	3	3	5	4	6	6	7	8	42
Coached teacher 4	2	3	5	3	6	7	5	8	39
Coached teacher 5	2	3	5	3	6	7	6	8	40
Total: Coached group	**11**	**15**	**25**	**19**	**30**	**34**	**31**	**40**	**205**

***Note*: Column headings are abbreviations for each of the effective Mathematics Teaching Practices chosen by the author. Conc => Build procedural fluency from conceptual understanding; Strug => Support productive struggle in learning mathematics; Goals => Establish mathematics goals to focus learning; St. Th => Elicit and use evidence of student thinking; Rep'n => Use and connect mathematics representations; Tasks=> Implement tasks that promote reasoning and problem solving; Disc => Facilitate meaningful mathematics discourse; Ques => Pose purposeful questions.**

The data for coached teachers in Table 4.3 reflects that each coached teacher's experiences were different from that of the other coached teachers. Therefore, the number of opportunities each teacher had to learn about each of the effective Mathematics Teaching Practices are sometimes different. For the coached teachers, the total includes learning opportunities from follow-up sessions *and* coaching sessions. The table provides evidence that coached teachers not only had more overall opportunities to learn, they also had more opportunities to learn about each individual effective Mathematics Teaching Practices compared to comparison teachers.

Figure 4.3 uses the totals for the group of comparison teachers and the totals for the group of coached teachers from the bolded rows in Table 4.3 to illustrate the distribution of teachers' opportunities to learn across the eight effective teaching practices for mathematics. The sets of bars illustrate the distribution of teachers' opportunities to learn about the effective Mathematics Teaching Practices in *Principles to Action: Ensuring Mathematical Success for All* (NCTM, 2014). The pairs of bars are ordered from the practice least encountered by the comparison group of teachers to the practices most encountered by the comparison teacher group, namely *facilitate meaningful mathematics discourse* and *pose purposeful questions*. Coached teachers' encounters with each practice are shown beside the number of encounters for comparison teachers. This display makes apparent that the number of opportunities to learn about each effective Mathematics Teaching Practice for coached teachers exceeded that of comparison teachers.

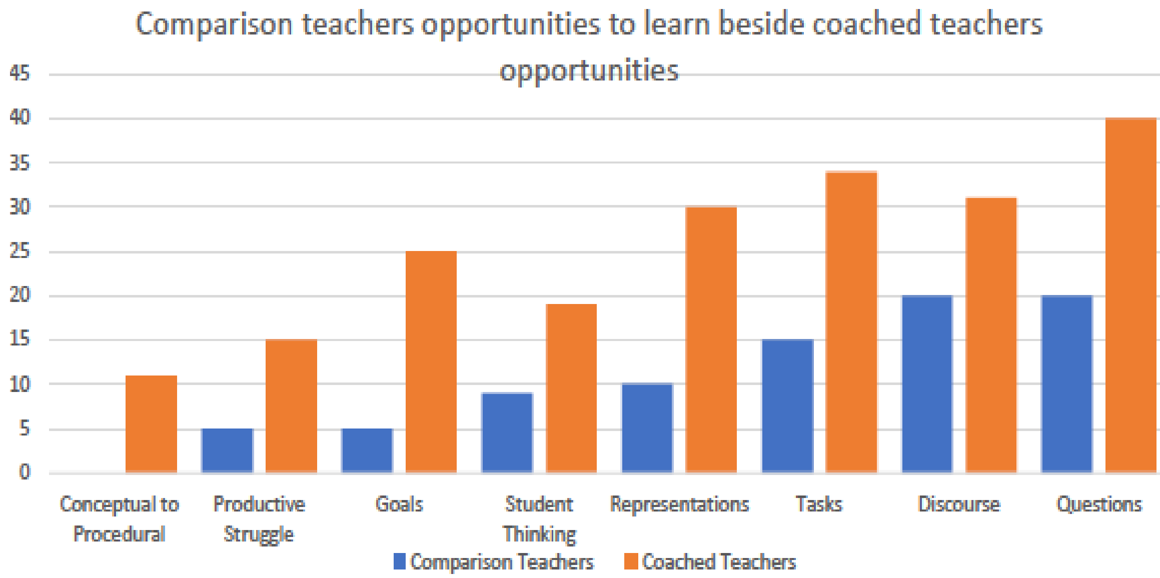


Figure 4.3 Comparison teachers' opportunities to learn about each practice compared to coached teachers' opportunities ordered from comparison teachers' least to most-encountered practice

Teachers' encounters with the eight effective Mathematics Teaching Practices elicit a set of side-by-side, stacked bars as well. Figure 4.4 exhibits the same data as Figure 4.3, except that each bar is stacked with the number of times an opportunity was rated as “some” on the bottom of the stack and the number of “extended” opportunities to learn on the top of the stack. Thus, Figure 4.4 shows not only the data gathered for each effective Mathematics Teaching Practice but also the data about the depth of the learning opportunities. Figure 4.4 makes it apparent that coached teachers had more overall experiences, more experiences with each effective Mathematics Teaching Practice, more experiences rated as “some” for every effective Mathematics Teaching Practice, and more experiences rated as “extended” for six of eight effective Mathematics Teaching Practices. The exceptions are *facilitate meaningful mathematical discourse* and *support productive struggle in learning mathematics.*

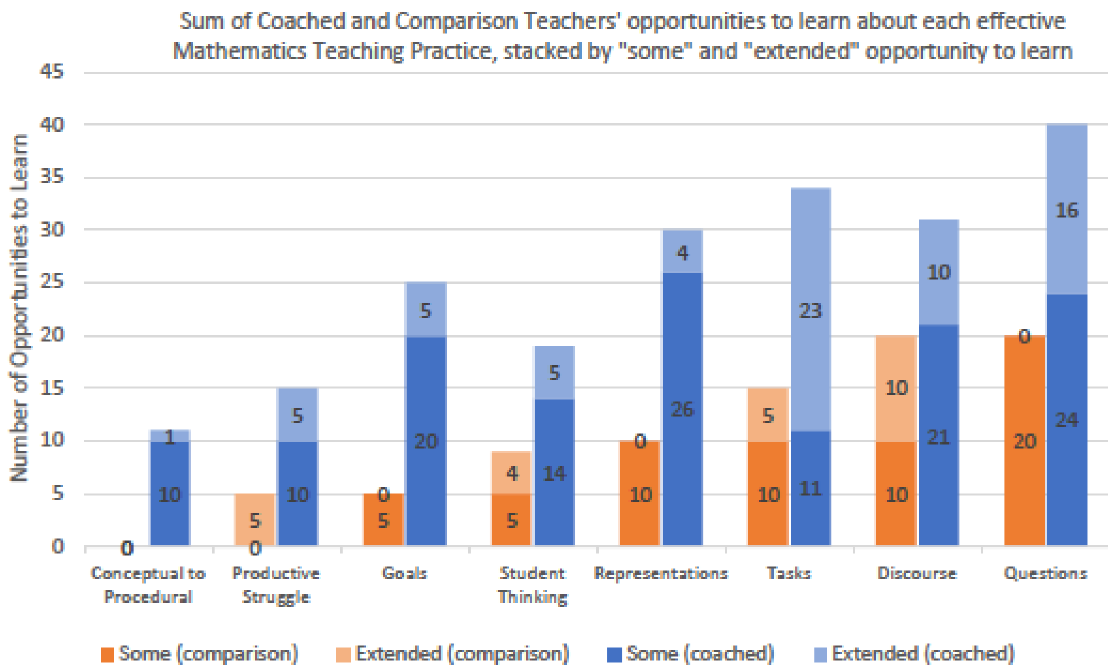


Figure 4.4 Opportunities to learn about effective Mathematics Teaching Practices shown by practice

Note: **Because one comparison teacher missed an afternoon of follow-up PD, the group of comparison teachers had only four extended opportunities to learn about student thinking instead of five. No coached teachers had an extended opportunity to learn about this practice during coaching. Although their group had one additional opportunity to learn about this practice, this occurs only because all coached teachers were in attendance at all the follow-up PD.**

In examining data concerning the individual effective Mathematics Teaching Practices, some stark difference in opportunities to learn emerge. Teachers in the comparison group had no opportunity to learn about the effective Mathematics Teaching Practice of *building procedural fluency from conceptual understanding* during the follow-up PD. Albeit the practice with which the coached teachers had the fewest encounters, each coached teacher had two or three chances to learn about *building procedural fluency from conceptual understanding.*

Another dramatic difference in opportunity to learn about a single effective Mathematics Teacher Practice is in the practice of *establishing goals to focus learning*. As a group, the coached teachers had five times as many chances to learn about goals as did the comparison group teachers. While all of the teachers in the study encountered the difference between learning goals and performance goals in one of the school year follow-ups, each of the coached teachers had four additional encounters with the practice of writing learning goals: one during each of the four coaching cycles. One of these four additional encounters with goals was an “extended” learning opportunity that was described earlier. (See section 4.1.1.2 of this chapter.)

Both coached and comparison teachers had multiple opportunities to learn about the effective mathematics teaching practice of *posing purposeful questions*. All the teachers participating in the study attended the four follow-up PD sessions where questioning was a portion of each day’s learning. Each of the coached teachers had four additional opportunities to learn about posing purposeful questions because each coaching cycle involved some discussion of this practice. Therefore, the coached teachers had twice the number of chances to learn about questioning, with each coached teacher having eight chances and each comparison teacher having four chances.

The greater difference in opportunities to learn about questioning is in the depth of the opportunities. While each of the follow-up PD sessions discussed questioning, there were no follow-up PD sessions that discussed questioning in-depth. In contrast, nearly every coaching cycle involved an in-depth opportunity to learn about posing purposeful questions. Four of the five coached teachers had three extended encounters to learn about questioning and one opportunity rated as “some” within their coaching cycles. The remaining coached teacher had four extended encounters with questioning–one in each of the coaching cycles. This accounts for the 16 extended opportunities for coached teachers to learn about *posing purposeful questions*, as seen in Figure 4.4.

Even though coached teachers had twice as many opportunities to learn about the practice of *posing purposeful questions*, it is actually the second lowest relative difference between the two groups. The only practice with a lower relative difference is the effective mathematics teaching practice of *facilitating meaningful mathematics discourse*. As a group, the five coached teachers had 31 opportunities to learn about discourse, while the group of five comparison teachers had one opportunity at each of the follow-up PD sessions for a total of 20 chances to learn about discourse, meaning coached teachers had 1.55 times as many opportunities to learn about discourse.

While *facilitating meaningful discourse* had the smallest relative difference between the coached and comparison teachers, two other practices had the smallest raw difference. The practices of *eliciting and using evidence of student thinking* and of *supporting productive struggle in learning mathematics* had a net difference of 10 occurrences when comparing the group of comparison teachers to the group of coached teachers. All the teachers in the study had one in-depth experience with productive struggle during a follow-up PD session. The coached teachers each had two additional chances to learn about productive struggle during their coaching cycles, but both of those learning chances were less intense than the one provided during follow-up PD. All teachers in the study had one in-depth experience rated as "extended" and one less comprehensive encounter with *eliciting and using student thinking* rated as "some" opportunity to learn. Each coached teacher had at least one more chance to learn about *eliciting and using student thinking* during their coaching cycles. As with the practice of *supporting productive struggle in learning mathematics*, the coaching experiences related to *eliciting and using student thinking* were not in-depth encounters.

Coached teachers had a total of 30 encounters with *using and connecting mathematical representations*: 26 (87%) rated as "some" and four (13%) rated as "extended." Comparison teachers had only ten encounters with *using and connecting mathematical representations*, all of which were rated as "some" opportunity to learn. The group of coached teachers had three times as many chances to learn about this practice compared to the group of uncoached teachers. This practice also saw the greatest raw difference between coached and comparison teachers' chances to learn about an effective Mathematics Teaching Practice. It is tied with *establish mathematics goals to focus learning* and *posing purposeful questions* for this distinction. All three practices have a difference of 20 between the groups of comparison and coached teachers.

With regard to the teaching practice of *implementing tasks that promote reasoning and problem solving*, coached teachers had 34 opportunities to learn, which is close to the number of encounters this group had with *using and connecting mathematical representations*, but the ratio of "some" to "extended" opportunities to learn is very different from the ratio for the aforementioned practice. Coached teachers had 11 opportunities rated as "some" (32%) and 23 "extended" opportunities (68%). Comparison teachers had 15 total chances to learn about tasks that promote reasoning and sense making: ten rated as "some" (67%) and five rated as "extended" (33%) So, the ratio of some to extended opportunities to learn for the comparison group of teachers is the inverse of the ratio for the coached teachers. A high proportion of coached teachers' interactions with learning about tasks were rated as "extended" because the task is the avenue through which the students will interact with the mathematics. The task frames the activity through which students will draw closer to attaining the learning goal (Stein et al., 1996), so it becomes the focus of the lesson, only to be preceded by becoming the focus of the planning and to be followed by becoming the focus of the lesson's debrief.

4.1.3.2 Qualitative results

Examining the qualitative side of the analysis of which effective Mathematics Teaching Practices were encountered affords a multitude of data. Table 4.4 provides brief examples of teachers' opportunities to learn about each practice. It is followed by expansion upon a few learning opportunities in narrative form within the text of this chapter. Within Table 4.4, there are two examples for each effective Mathematics Teaching Practice pertaining to all teachers in the study: one example of an opportunity to learn rated as “some” and one example of a learning opportunity rated as “extended.” Two additional examples pertaining only to coached teachers are likewise provided.

Narrative examples for encounters with many of the effective Mathematics Teaching Practices have been previously provided; however, illustrations for effective Mathematics Teaching Practices not previously put into narrative form may further elucidate the qualitative differences between encounters that occurred in the follow-up PD sessions and encounters from coaching. One practice that has not yet been discussed in the narrative is *supporting productive struggle in learning mathematics*. All teachers had an “extended” opportunity to learn about *supporting productive struggle* in the third follow-up PD session, where the theme for a portion of the day was growth mindset (Dweck, 2006). The session began with a discussion of four “research findings with implications for learning” (Stanford Graduate School of Education, n.d.). Specifically, those finding are:

- Every child can learn at high levels.
- Mistakes grow your brain.
- Messaging from adults influences children’s achievement and
- When you believe in yourself your brain operates differently.

Table 4.4 Examples of opportunities to learn about each effective Mathematics Teaching Practice

	Common experiences for coached & comparison teachers		Experiences unique to coached teachers	
	Some	Extended	Some	Extended
Establish Mathematics Goals to Focus Learning	Discussed learning goals within the context of a reading.	Not applicable	Clarified difference between learning and performance goals. Came to consensus on learning goal(s).	Engaged in an iterative cycle of choosing and revising or refining learning goal(s).
Implement tasks that promote reasoning and problem solving	Revisited the *Task Analysis Guide* (Stein, et al., 2000), examining same content addressed at each level of cognitive demand.	Examined Benchmark Task Grid (Boston, 2012d), rated sample tasks, described characteristics, introduced *Task Analysis Guide* (Stein, et al., 2000), examined research, surfaced previous tasks, etc.	Discussed whether task was high level as written and as implemented. Discussed challenges and affordances of the task.	Used *Task Analysis Guide* (Stein et al., 2000) for discussion of whether the planned task was high level. Jointly implemented task. Debriefed and considered a future task. Coach shared other resources.
Use and connect mathematical representa-tions	Discussed representations students may use during a number talk and how to connect the representations therein.	Not applicable	Discussed the lack of representations directly elicited by the chosen task and how to remedy that situation.	Discussed representations students might use, relative strengths of each, struggles students might encounter with each. This led to other helpful representations and models. Discussion continued in debrief.
Facilitate meaningful mathematical discourse	Engaged in the *5 Practices for Orchestrating Productive Mathematics Discussions* (M. S. Smith & Stein, 2011).	Discussed facilitating discourse in context of a case; connected to *Principles to Action* (NCTM, 2014) Discussed and used the *5 Practices* (M. S. Smith & Stein, 2011).	Discussed strategies for subtly quieting dominant voices to balance student group participation and allow quieter voices to be heard.	Not applicable

	Common experiences for coached & comparison teachers		Experiences unique to coached teachers	
	Some	Extended	Some	Extended
Pose purposeful questions	Discussed appropriate and not-as-appropriate questions to avoid funneling student thinking.	Not applicable	Discussed focusing versus funneling questions in relation to the planned and employed questions for the lesson.	Planned and revised assessing and advancing questions; Teacher observed coach using the questions in the monitoring phase and other connecting questions in the share/discuss phase. Discussed how this focused thinking.
Build procedural fluency from conceptual understand-ing	Not applicable	Not applicable	Discussed where the planned task was on the progression from developing conceptual understanding to finding more generalized methods to eventual procedural fluency.	Discussed how and why the planned series of connected tasks would help students work towards procedural fluency.
Support productive struggle in learning mathematics	Not applicable	Read article about mindset and discussed findings related to growth versus fixed mindset (Dweck, 2006).	Discussed how and why the chosen and implemented task did not support productive struggle.	Not applicable
Elicit and use evidence of student thinking	Discussed how student talk provides evidence of student thought processes as well as student level of proficiency.	Engaged in magnetic quote activity. Watched videos. Discussed research brief and encountered Formative Assessment Lessons (FALs).	Discussed how the task and its facilitation would elicit student thinking via questions, written work, and manipulative models. Revisited in the debrief with instances from the lesson.	Not applicable

The last finding concerning "belief in self" led to more extensive discussion of growth versus fixed mindset (Dweck, 2006), how and when to best use praise, and how and when not to use praise. Next, participants read short articles about mindset (Dweck, 2007; Mindset Works Inc., 2016) as it related to supporting students in productively struggling. Then, participants viewed online videos, visited websites, and examined teacher-created posters to support the development of growth mindset and support students' willingness to engage in struggle in mathematics classrooms. To conclude the portion of the day about mindset, teachers discussed how best to communicate ideas about mindset and struggle to parents, administrators, and other teachers.

A coaching cycle illustrating the use of productive struggle that was rated as "some" opportunity to learn occurred during the last coaching cycle of the year. For that cycle, one of the coached teachers chose a task she believed would provide an indication of her students' progress towards understanding addition of two-digit numbers. The task, called *Ford and Logan add 45+36* (Illustrative Mathematics, 2016a) (See Appendix K.3), asked students to solve 45+36 and then examine and analyze two other fictitious students' methods for solving the addition problem. During the planning portion of the coaching cycle, the teacher and coach discussed scaffolding student thinking without taking over via the use of questions geared specifically towards students' anticipated struggles. The pair discussed providing encouragement for students to reflect on their own strategies as well as the strategies of Ford and Logan from the task. The pair also discussed the best ways to acknowledge student contributions, especially during the share-and-discuss phase of the lesson. During task implementation, the teacher assured that adequate time was provided for students to struggle with the task.

According to *Taking Action: Implementing Effective Mathematics Teaching Practices in K-Grade 5* (Huinker & Bill, 2017), each of the moves mentioned above supports students' productive struggle in mathematics. In this coaching cycle, the actions were personalized by the teacher to her class for that particular lesson. For example, during the implementation of this task, student groups were struggling with making sense of one of the fictitious character's ways of solving the addition problem. The coach-teacher team had anticipated this and planned specific questions to support students in making sense of the solution strategy without removing the struggle. This teacher's opportunity to learn about *supporting productive struggle* was rated as "some" because it was not the main focus for this coaching session and was not a lengthy encounter with the practice of *supporting productive struggle*. However, this teacher's encounter with the practice was qualitatively different from that provided during the follow-up session on growth mindset (Dweck, 2006), when the learning opportunity was rated as "extended."

Just as all teachers had an opportunity to learn about *supporting productive struggle* during one of the follow-up PD sessions, all teachers had an opportunity to learn about the effective Mathematics Teaching Practice of *using and connecting representations* during the second follow-up PD session. The main body of this session dealt with *facilitating meaningful mathematical discourse* and *implementing tasks*. To start the session, participants engaged in the *Hungry Caterpillar Task* and read the *Case of Ms. Bouchard* (Huinker & Schefelker, 2016). *Using and Connecting representations* entered the session when the large group discussed (1) Ms. Bouchard's use of different pieces of student work that had drawings or number sentences on them and (2) how Ms. Bouchard helped students connect one representation to another during the share-and-discuss portion of the case. This learning opportunity was rated as "some."

During coaching, an instance of *using and connecting representations* occurred when the teacher and coach collaborated to implement a variation of the *Building a Rabbit Pen Task* (Math Design Collaborative, 2015). (See Appendix K.4.) Coach and teacher conferred about the learning goal for this lesson, created monitoring charts for the task, and began to talk in detail about the student solution strategies. At this point, the coach-teacher team realized that students would likely have trouble creating a drawing of the rabbit pen on the grid provided with the task, so they brainstormed available tools students might use to create a concrete model of the pens. Popsicle sticks were the tool of choice and provided a way for students to model and manipulate the pretend rabbit pens during the lesson. In this example, the rating for the practice of *using and connecting representations* was “some.” Even though the discussion of representations was not rated as an “extended” opportunity to learn about this teaching practice, the representations were an important part of students’ engagement in the task.

In summary, as expected, coached teachers in the study had more opportunities to learn about the effective Mathematics Teaching Practices than the comparison teachers. Coached teachers experienced four coaching cycles, each of which entailed some learning about at least four of the eight effective teaching practices for mathematics, in addition to the four follow-up PD sessions, each of which involved learning about four or five of the practices. All told, the coached teachers had more than two and a half times as many chances to learn about effective practices associated with ambitious mathematics teaching (204) compared to comparison teachers (84). The coached teachers had more than twice as many experiences where they learned “some” about a practice with 136 chances compared to 60. They had nearly three times as many in-depth chances to learn about an effective mathematics teaching practice with 69 chances compared to 24. Additionally, the encounters with the effective Mathematics Teaching Practices were qualitatively different for coached teachers than for the comparison group teachers who only attended PD sessions. While coaching added quite a bit to teachers’ opportunities to learn about ambitious mathematics teaching, the following section will discuss the differences in teachers' *use* of ambitious teaching practices.

4.2 Use of ambitious teaching practices

Recall from the chapter introduction that the second question this study seeks to address is: *What is the impact on teachers'* use *of ambitious teaching practices when content-focused coaching is added to professional development?* The Instructional Quality Assessment (IQA) (Boston, 2012) (Appendix H) was used to measure teachers' use of ambitious teaching practices[11]. The study author and an outside evaluator[12] used four of the IQA Academic Rigor (AR) rubrics pre- and post-coaching activities to assess the ten teachers' use of ambitious teaching practices. Before the start of any coaching activities, the two evaluators performed consensus scoring for each of the coached and comparison teachers in the study on one lesson. After the conclusion of all coaching activities, the two evaluators again performed consensus scoring for each of the five coached and five comparison teachers. The four rubrics from the IQA used in this study were AR1: Potential of the Task, AR2: Implementation of the Task, AR3: Student Discussion following Task, and AR-Q: Rigor of Teachers' Questions[13][14][15]. This section of the results chapter is spent

---

[11] Recall that while the study was initially to use both the IQA and the Effective Teacher Practices checklist, use of the checklist had to be abandoned because it was never validated before the conclusion of the study.

[12] Special thanks to Dante Orsini for serving as the outside evaluator for this study and to Dr. Melissa Boston for her role facilitating the work.

[13] Classroom observation rubrics were employed. Assignments collection rubrics were not employed in order not to burden participating teachers.

[14] While ambitious mathematics instruction is aligned with both the AR and AT rubrics, it was recommended by the dissertation committee that AR rubrics be used in observing ambitious teaching practices over AT rubrics. It was impractical to use the full set of rubrics for classroom observation.

[15] The fifth AR rubric, AR-X (Residue), was not used. It was still in pilot form at the time of this study.

analyzing the IQA scores. The composite scores from the IQA are analyzed in subsection 4.2.1. Following that, in section 4.2.2, the data gathered from individual rubric scores is analyzed. This portion of the chapter concludes with a section on qualitative results.

4.2.1 IQA composite scores

Table 4.5 provides the data from the consensus scoring on the IQA AR rubrics used in this study. Because four AR rubrics, scored from 0 to 4, were employed, the maximum composite score is 16 for this study. Two of the comparison teachers increased their composite IQA scores. One comparison group teacher, the highest scoring teacher in the comparison group, had the same composite score in the fall and in the spring. Two comparison teachers' composite scores decreased from Fall 2017 to Spring 2018; one by three points and the other by five points. In comparison, four of the five coached teachers increased their composite scores on the IQA from Fall 2017 to Spring 2018 with one teacher increasing the composite score by seven points, which is the greatest change in IQA composite score seen in this study. The only coached teacher who did not increase in composite IQA score decreased by one point.

Table 4.5 Summary of IQA AR rubric scores for Fall 2017 and Spring 2018

	Composite IQA(Fall'17)	AR 1	AR 2	AR 3	AR-Q	Composite IQA(Spr18)	AR 1	AR 2	AR 3	AR-Q
Comparison Teacher 1	11	3	3	1	4	8	2	2	2	2
Comparison Teacher 2	13	4	3	3	3	8	2	2	2	2
Comparison Teacher 3	12	3	3	3	3	14	4	3	3	4
Comparison Teacher 4	14	4	4	2	4	14	3	4	3	4
Comparison Teacher 5	6	2	2	1	1	12	4	2	2	4
Coached Teacher 1	7	2	2	2	1	14	3	4	3	4
Coached Teacher 2	10	3	3	2	2	14	3	4	3	4
Coached Teacher 3	11	2	2	3	4	15	3	4	4	4
Coached Teacher 4	8	3	1	2	2	7	2	2	1	2
Coached Teacher 5	7	2	2	1	2	10	2	2	2	4

The Mann-Whitney nonparametric t-test for independent samples was used on each teacher's composite IQA score to analyze differences between the comparison teachers and the coached teachers. The Mann-Whitney test uses rankings instead of raw scores to focus attention on the ordered relationships instead of the spread of the data. At the onset of the study, comparison teachers' rankings were better than coached teachers, with comparison teachers having four of the top five rankings ($T_{initial}$(comparison) = 20.5; $T_{initial}$(coached) = 34.5), but the difference between the groups was not significant ($z_{initial}$ = -1.36, with area beyond z = 0.0869). At the conclusion, coached teachers' rankings were better than comparison teachers, with coached teachers having three of the top five rankings. Four of the five coached teachers increased their overall rank (T_{final}(comparison) = 30; T_{final}(coached) = 25). Although the difference between the groups was directionally different, the difference was still not significant (z_{final} = 0.42; with area beyond z =

0.3372). Thus, findings from the IQA composite scores for the four AR rubrics show no significant differences in instructional practice between the two groups of teachers either before or after coaching during the 2017-18 school year.

Additionally, the group of coached teachers did not significantly improve their composite IQA scores from Fall 2017 to Spring 2018. In comparing coached teachers' composite IQA scores from fall to spring, four of five teachers improved their composite IQA score. The test of significance for this data is the Wilcoxan Signed-Rank test. This test is for nonparametric data with matched samples, which applies to the coached teachers' scores for fall and for spring. By using ranks instead of the actual composite IQA scores, the Wilcoxan Signed-Rank test removes concerns with the amount of variation within the samples. Unfortunately, with a sample size of five, because one of the five coached teachers did not improve their composite IQA score from Fall 2017 to Spring 2018, the result for composite IQA scores was not significant[16] (n=5, W=13). To be significant with only five subjects, the sum of the signed ranks (W) must be 15, which is the maximum value W can have for n=5. Thus, findings from the IQA composite scores for the four AR rubrics show no significant differences in instructional practice for the coached teachers from Fall 2017 to Spring 2018.

Continuing to examine the composite IQA scores yields other information. The individual AR rubrics are scored from a low score of 0 to a high score of 4. A high rating on an individual AR rubric is considered a score of 3 or 4 (Boston, 2012a). To earn what is considered a "high" composite IQA score, the teacher cannot have all the individual AR rubric scores at 2 or lower. Said another way, the teacher has to earn at least one 3 or higher on an AR rubric, even if all the remaining scores are considered low-level ratings. Thus, "high" composite IQA scores must total

---

[16] No z score is calculated when n is small because it is feasible to list all possible cases.

9 or greater for the set of four scores. Only two of the coached teachers had a high composite score before coaching began, while four of the comparison group teachers had a high composite score in Fall 2017. After the conclusion of coaching, four of the coached teachers earned a high composite score. Three of the comparison group teachers had a high composite score in Spring 2018. See Tables 4.6. and 4.7 for this data.

Table 4.6 Number of teachers in each group earning low (< 9) or high (≥ 9) composite IQA scores

	Composite IQA–Fall 2017		Composite IQA–Spring 2018	
	0 ≤ sum <9	9 ≤ sum ≤ 16	0 ≤ sum < 9	9 ≤ sum ≤ 16
Comparison teachers	1	4	2	3
Coached teachers	3	2	1	4
Total	4	6	3	7

Table 4.7. Number of coached teachers earning low (<9) or high (≥9) composite IQA scores in fall and spring

	Composite IQA – Coached Teachers		
	0 ≤sum<9	9 ≤sum ≤16	Total
Fall 2017	3	2	5
Spring 2018	1	4	5
Total	4	6	10

Because the expected values for the cells in these tables are less than five, chi squared cannot be used. Instead, the Fisher Exact Probability test was used to examine the difference between the coached and comparison teachers in Spring 2018. One can likely determine from inspection that these results do not show a significant difference between coached and comparison teachers ($p = 0.50$). The Fisher Exact Probability test was again used to examine the difference between the number of high composite scores for coached teachers in Fall 2017 and Spring 2018.

Results show that this difference is also not significant (p = 0.26). Examining the findings for the IQA composite scores shows no significant differences between the group of comparison teachers and the group of coached teachers. Additionally, findings from the IQA composite scores show no significant differences in instructional practice for the coached teachers from Fall 2017 to Spring 2018.

4.2.2 Academic Rigor rubric scores

The coached teachers, taken as a group, exhibited significantly higher scores on the *individual* AR rubrics, in contrast to findings for the *composite* IQA scores. See Table 4.5 for these scores. In comparing coached teachers' IQA AR rubric scores from the fall to the spring, fully 18 of 20 individual AR ratings for the five coached teachers taken from AR1, AR2, AR3, and AR-Q either improved or were maintained. Of those 18 improved or stable AR ratings, 16 belong to the four teachers who improved their composite IQA scores. In other words, four of the five coached teachers either improved or maintained *every* one of the AR ratings that comprise their composite IQA score. In fact, 12 of the 16 AR ratings for the four improved teachers increased from fall to spring; four of these 12 improved ratings increased by two points; and one rating increased by three points.

This three point increase from a "1" to a "4" is seen on AR-Q and pertains to a coached teacher who increased the score on every one of the four AR rubrics used in the study. Two other coached teachers increased the scores on three of the four AR rubrics while maintaining their score on the fourth AR rubric. Another coached teacher maintained scores on two of the four rubrics and increased scores on the other two rubrics. The fifth coached teacher had a different profile. The

AR rubric scores for this teacher increased on one rubric, stayed the same on a second rubric, and decreased on the remaining two rubrics. See Figure 4.5 for a visual display of this data.

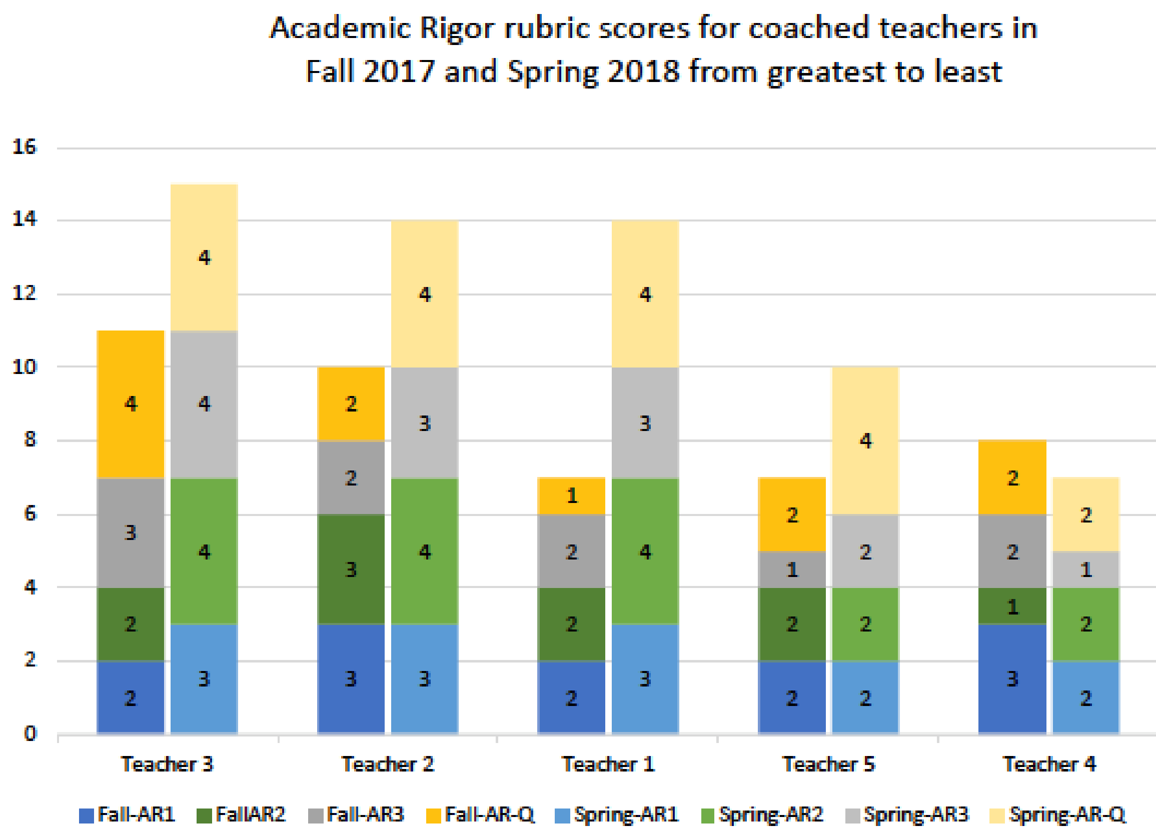


Figure 4.5 Stacked side-by-side bar graph with individual AR rubric scores for coached teachers

As with the composite IQA scores, the test of significance is the Wilcoxan Signed Rank test for nonparametric data with matched samples. The difference in the test for the AR rubric scores is that the sample size is no longer a small one. There are now 20 rubric scores to compare (4 AR rubrics x 5 coached teachers) instead of five composite IQA scores to compare. Because so many of the 20 individual AR rubric scores for the coached teachers increased or remained the same and only two of the 20 decreased, the results of the Wicoxan Signed Rank test for this data were significant[17] (n = 15; W = 98; z = 2.77, P = 0.0028).

Continuing to examine individual AR rubrics yields more information. As noted previously, the AR rubrics are scored from 0 to 4, and a high rating on an individual AR rubric is a score of 3 or 4 (Boston, 2012a). At the onset of the study, the group of five comparison teachers earned more high ratings on the AR rubrics than did the group of five coached teachers. There were 14 high ratings for the comparison group as opposed to five high ratings for the coached group. At the conclusion of the study, coached teachers earned more high ratings on the AR rubrics than comparison teachers: 13 high ratings for coached teachers and ten for comparison teachers. The chi squared test determined that coached teachers had earned more high scores than expected in Spring 2018, and comparison teachers had earned fewer high scores than expected. Thus, the chi squared test shows a significant difference between the groups ($\chi^2 = 2.75$, df = 1, P = 0.049) for the number of AR rubrics earning high scores.

In general, larger grain results from the IQA, such as the composite scores, show a lack of significant difference between the coached and comparison teachers in Spring 2018, when it was hoped coached teacher would score higher. The composite IQA scores also show a lack of significant change in scores from fall to spring for coached teachers. However, results are significant when looking at finer grain results, such as the individual AR rubric scores. Coached teachers made significant gains in their AR rubric scores from Fall 2017 to Spring 2018. Coached teachers also earned significantly more high ratings on the AR rubrics than did comparison teachers in the Spring 2018. Aside from the quantitative differences in coached teachers' IQA AR

[17] There were 20 individual AR rubric scores among the five coached teachers in Fall 2017 and Spring 2018. Five of these AR rubric scores stayed the same from fall to spring, so n=20-5=15 for the Wilcoxan Signed Rank test.

rubric scores, there existed qualitative changes in coached teachers from before coaching to afterwards. This dissertation now turns to examining some of the qualitative changes.

4.2.3 Qualitative analysis

Although the overall IQA composite scores do not show a significant increase in coached teacher scores, the changes in their composite scores provide some qualitative information about differences in ambitious teaching practices. The individual AR rubric scores provide additional qualitative data. As Boston wrote, "Score levels on the IQA rubrics enable quantitative and qualitative interpretations, as scores represent different levels of instructional quality and specific features of mathematics instruction" (2012a, p. 95). This section will examine qualitative changes, some dramatic and some more subtly, in some of the coached teachers' practices as evidenced by the indicators and descriptors for the set of four AR rubrics employed in this study. Following that, the chapter contrasts changes in coached teachers' practice with qualitative changes in some of the comparison teachers' practice.

Within both groups of teachers, there were changes in the observed instruction. The most drastic improvement occurred with the first coached teacher. Coached Teacher 1, a first grade teacher, earned improved AR scores on each of the four rubrics used in the study and improved the overall IQA score by seven points. Her classroom was qualitatively different from the pre-coaching observation to the post-coaching observation. This teacher initially chose a low-level task during the pre-coaching observation, but used a high-level task after coaching. According to the AR1 rubric, this shows the teacher used a task that had "the potential to engage students in complex thinking or in creating meaning for mathematical concepts, procedures, and/or relationships" (Boston, 2012b, p. 9). She moved from having her students engage in performing a procedure to allowing them to explore mathematical concepts. This corresponds to a change in score on AR2 from a 2 to a 4. The discussion after the task moved from a recitation of steps to solve, to a sharing of different strategies; a shift on AR3 from a score of 2 to a score of 3. The largest change for this teacher occurred with AR-Q. The score changed from low, where the teacher asked only procedural questions, to the highest rating. During the post-observation, the teacher asked questions that allowed students the chance to explain their thinking and elaborate on their written work by verbalizing it. These changes in the type of task used and the manner in which it was implemented, including the questions asked and the discourse that occurred during the monitoring and share-and-discuss phases of the lesson account for a qualitatively different teaching experience for the teacher and a qualitatively different learning experience for the students in this classroom when comparing the pre-coaching observation to the post-coaching observation.

Other qualitative changes are seen with examining changes in teacher scores on some of the individual AR rubrics employed. Some large qualitative changes occurred with the rubric examining teachers' questioning practices, AR-Q. Aside from Coached Teacher 1, two other coached teachers improved their AR-Q rubric scores from a low rating to the highest rating. To earn this rating, teachers must "consistently ask academically relevant questions" (Boston, 2012b, p. 39). These types of questions allow students access to the underlying mathematical ideas of the lesson and allow connections to be made by students. Hearing the coached teachers ask mathematically important probing or focusing questions; seeing the teachers encourage student reflection on the mathematics; witnessing the teachers' use of questions to help make the mathematics visible to students; and hearing the students discuss their reasoning with classmates as a consequence of the teachers' questions were qualitative changes seen in the classrooms of the teachers whose AR-Q scores changed from low to high. These changes indicated the use of ambitious teaching practices in classrooms where they had not previously been used. These teachers were now engaging in the effective Mathematics Teaching Practices of *posing purposeful questions* and *facilitating meaningful mathematics discourse* (NCTM, 2014). Of the four coached teachers who used questioning tactics aligned to the highest rating on the AR-Q rubric during the post-coaching observation, three had not done so before coaching began.

Coached teachers also made improvements on the AR2 rubric for Implementation of the Task. Before coaching began, four of the five coached teachers either engaged students in a procedural exercise with no connections to the meaning of the procedure or the underlying mathematics, or the teacher engaged students in a memorization activity. One teacher engaged students in a procedures with connections task, rated as “3” for Implementation of the Task. However, at the conclusion of coaching, three of the five coached teachers had ratings the highest level, a “4” on the AR2 rubric. This means that the students in their class were “engaged in exploring and understanding the nature of mathematical concepts, procedures, and/or relationships” (Boston, 2012b, p.21). The students explored connections among the mathematical procedure and the underlying meaning or engaged in problems where they, themselves had to uncover the mathematics to solve the problem. For the two teachers whose initial AR2 scores were low, the change to the highest rating indicated a qualitative shift from students reproducing facts or procedures to students engaging in doing mathematics. As with AR-Q, this indicates teachers’ capabilities to use ambitious teaching practices. Higher ratings on the AR2 rubric indicate use of the effective Mathematics Teaching Practice of *implementing tasks that promote reasoning and sense making.* Additionally, implementation of tasks at a higher level on the AR2 has the potential to build the conceptual understanding on which future procedural fluency is built (NCTM, 2014) for students in these coached teachers' classes.

Aside from the first coached teacher mentioned in this section of the chapter, two other coached teachers also earned all high-level ratings on the post-observation with the IQA. One of these teachers ended with two "4" ratings, and the other ended with three "4" ratings. Earning more high-level ratings, and especially earning the highest ratings, on the AR rubrics generally indicates more ambitious mathematics instruction, since "(t)he IQA assesses elements of ambitious instruction in mathematics; specifically, the level of instructional tasks and task implementation, opportunities for mathematical discourse, and teachers' expectations" (Boston, 2012a, p. 76). While this study did not employ the entire IQA tool, the four chosen AR rubrics still provide indicators of (1) the use of challenging tasks that promote mathematical reasoning, encourage the development of conceptual understanding, and allow for the use of varying mathematical representations; (2) the employment of purposeful questions that can form the basis for meaningful mathematical discourse; (3) opportunities for students to engage in complex thinking and productive struggle and subsequently demonstrate evidence of this to their teacher. Thus, the chosen rubrics provide connected qualitative and quantitative evidence of ambitious teaching practices.

The group of five comparison teachers' AR scores ended with a different profile from the coached teachers. While two AR scores decreased for the coached teachers, eight AR rubric scores decreased in the comparison teacher group. One of the comparison teachers began with all high scores of 3 or 4 on the four IQA AR rubrics but had all low scores of 2 at the conclusion. This decrease was the most drastic decline in teacher scores. Another comparison teacher whose AR rubric scores declined had three of four high scores at the pre-coaching observation but had all low scores for the post-coaching observation. According to Boston (2012a), lower ratings on the rubrics mean "students are rarely asked to describe, explain, justify, prove, or generalize their mathematical work and ideas by instructional tasks, and rarely provide complete and thorough descriptions, justifications, proofs, or generalizations during classroom instruction" (p. 94). Thus, the comparison teachers whose scores dropped did not engage in ambitious teaching practices during the post-coaching observation, which means that students' opportunities to deeply learn and the likelihood of students' development of thorough understanding of mathematical ideas were diminished.

4.3 Conclusion

It is clear from the quantitative data that coached teachers had many more opportunities to learn about the effective Mathematics Teaching Practices over the course of this study when compared to comparison group teachers, who attended follow-up PD sessions but received no coaching. The increased opportunities included more "extended" opportunities *and* more opportunities that were not as in-depth. Additionally, the qualitative data demonstrates how coached teachers' opportunities to learn about effective Mathematics Teaching Practices were different than comparison teachers'. Coached teachers' opportunities to learn also had direct and immediate use with students because the planned lessons were enacted on the same day as the planning conference for the coaching cycle occurred.

This study showed mixed results as to whether coached teachers used more ambitious teaching practices in their classroom lessons following coaching. Examining the larger grain data does not show significant differences between coached and comparison teachers, as evidenced by

IQA composite scores. However, in examining smaller grain information, there exist improvements in teachers' AR rubric scores within the composite IQA. Some coached teachers made large improvements in their AR rubric scores (and the composite IQA score) from before coaching to after, but the anecdotal and qualitative data gathered before, after, and throughout the coaching cycles is telling as well. The teachers who made the greater strides in scores and ratings were also the teachers who found and used cognitively demanding tasks or a series of such tasks, which is reflected in AR1. They were the teachers who continued to ask probing questions and generate more discussion about the mathematics as reflected in AR-Q and AR3. This shows movement towards more ambitious teaching practices in the coached teachers' classrooms. However, movement towards ambitious instruction on the part of the coached teachers is especially seen in maintenance of cognitive demand upon implementation of the task, as reflected in the AR2. It has previously been shown in research that maintaining the cognitive demands of mathematical tasks has substantial impact on student learning. (Boston & Smith, 2011; Stein et al., 2009) In fact, the largest increases in student learning have been shown to occur when students interact with high-level, cognitively demanding mathematical tasks on a regular basis (Boaler & Staples, 2008; NCES, 2003; Stein & Lane, 1996). Thus, it is possible that for the coached mathematics teachers participating in this study, the proximal, in situ professional development in the form of content-focused coaching paired with outside-the-classroom professional development facilitated a change in their pedagogical practices from traditional to more ambitious in nature.

5.0 Discussion

The purpose of this study was to determine whether content-focused coaching that is paired with outside-the-classroom professional development facilitates a change in classroom practice towards more ambitious teaching for elementary mathematics teachers. This chapter begins by summarizing the results and sharing the conclusions of the study regarding the addition of content-focused coaching to outside-the-classroom professional development. Then, possible explanations for the results of the study are offered. Following these explanations, the chapter turns to contextualizing the findings of the study. Next, the chapter examines possible implications of the study's findings to discuss what may be learned from this investigation and how the study's findings can inform professional development efforts involving coaching. The chapter then turns to a discussion of some limitations of the study, and concludes with recommendations for future research.

5.1 Conclusions

5.1.1 Summary of results

One conclusion drawn from this study is clear. In response to the first research question, coached teachers received more and different exposures to and experiences with the effective Mathematics Teaching Practices that represent ambitious mathematics instruction for this study. When content-focused coaching was added to outside-the-classroom professional development, coached teachers had significantly more *opportunities to learn about* ambitious teaching practices. Results of this study showed that coached teacher had more opportunities than comparison teachers to learn about the effective teaching practices for mathematics, viewed through several different lenses. Coached teachers had more overall opportunities to learn about the effective Mathematics Teaching Practices; in fact, coached teachers had 2½ times as many opportunities to learn about these practices compared to uncoached teachers. Coached teachers also had more opportunities to learn that were in-depth, and coached teachers had more opportunities to learn about each of the eight individual effective Mathematics Teaching Practices.

The response to the second research question is not as clear as the response to the first. The question asks about the impact on teachers' *use of* ambitious teaching practices when content-focused coaching is added to professional development. The answer to the question depends upon the grain size used to quantitatively analyze the data. When examining the larger grained results provided by the composite IQA scores, coached teachers did not use significantly more ambitious teaching practices than control teachers, and coached teachers did not change their teaching practices to make them more ambitious from the beginning to the end of the study. In contrast, when examining smaller grained results provided by the AR rubric scores from the IQA, the results are different. Coached teachers significantly improved the scores on the AR rubrics from fall to spring, and at the end of the study they earned more high scores (3 or 4) on these AR rubrics compared to the uncoached teachers. Thus, the answer to the second research question about usage of ambitious teaching practices is not a simple "yes" or "no." Instead, it depends on whether one examines large-grained or small-grained results. Qualitative data aligned with the descriptors for the scores on the AR rubrics indicate that most coached teachers made improvements, sometimes dramatic improvements, to their teaching practices.

5.1.2 Explanation of the results

The results for the first research question aligned largely with expected results. As anticipated, the coached teachers received more exposures to the effective teaching practices for mathematics. One unanticipated result with respect to the opportunities to learn about the effective teaching practices, was in the depth of opportunities to learn about the individual practices. It was anticipated that the coached teachers would have a larger number of "extended" exposures to learn about every one of the eight effective Mathematics Teaching Practices. Even though the coached teaches had more "extended" exposures to six of the eight teaching practices, coached and comparison teachers had the same number of "extended" exposures for two of the practices. Namely, coached teachers did not have a greater number of "extended" opportunities to learn about *supporting productive struggle* or *facilitating meaningful mathematical discourse*.

Reflecting upon the nature of the coaching in this study leads one to conclude that it should not have been anticipated that the coached teachers would necessarily have more "extended" opportunities to learn about every one of the effective Mathematics Teaching Practices. This is because of the adaptive nature of the coaching activities in this study. After the co-planning session for the initial coaching session, wherein the coach had pre-planned a very similar conversation about learning goals with each of the coached teachers, the subsequent coaching activities were largely determined based on individual coached teachers' grade level, curriculum, stated strengths and needs, and perceived strengths and needs. Thus, each coached teacher's experiences departed from the others' experiences. Examining this in retrospect, it should not have been expected that all of the individual effective teaching practices for mathematics would necessarily receive more in-depth exposure for the coached teachers compared to the uncoached teachers, especially with only four coaching sessions for each coached teacher in the scope of the study.

The results for the second research question were partially aligned with expected results. An unexpected result for this study was that composite IQA scores for the coached teachers did not significantly increase from before the start of the study to after the conclusion of the study. The Wilcoxan Signed-Rank test of significance was used when comparing coached teachers composite IQA scores from before coaching to after coaching. This test is used for nonparametric data with matched samples. The test uses ranks instead of absolute scores, so the order, along with whether scores increased or decreased, is important. There is no concern about whether the scale of measurement (composite IQA scores in this case) is an equal-interval scale. For this data set, there are only five sets of scores to compare, so only five rankings (from 1 through 5) were generated, based on the relative magnitude of change. Then, signs are attached to the rankings, according to whether there was an increase or decrease in score. N is less than 10 for this study, so the sample size is small. The maximum sum of the rankings (**W**) is 15 (1+2+3+4+5). To be significant, the sum of signed rankings in this case has to be 15. That means all the IQA scores had to increase for all the signs of the rankings to be positive. However, one of the five coached teacher's IQA score decreased, which means one of the signs was negative. (The score decreased by one total point, but the amount of decrease does not matter with the Wilcoxan Signed-Rank test.) The **W** was 13 (-1+2+3.5+3.5+5). If the teacher whose score decreased had increased the composite IQA from fall to spring by any amount then the Wilcoxan Signed-Rank test would have given a significant result. Results would also have been significant if the teacher's composite IQA score had remained the same. In that case, the maximum **W** would have changed to 1+2+3+4=10, and the sum of signed ranks (W) would have been 10 (1+2.5+2.5+4). All coached teachers' IQA score did not stay the same or increase, which means the results for composite IQA scores are not significant.

A possible reason why the change in coached teachers' composite IQA scores from before coaching to after coaching was not significant involves the length of the study and the role of the coach. Previous studies of coaching show there is a transition period when coaches are learning about, adjusting to, and developing in their new role (Campbell, 2012; Chval et al., 2010). Previous studies also show that the effect of coaching on student achievement begins to appear during the second year of implementation and continues thereafter (Campbell, 2012; Matsumura et al., 2013). This study occurred over only one school year with only four coaching sessions per teacher. Thus, there may not have been adequate time for teachers to adjust to the expectations of coaching or for the coach to develop a strong identity in her role. The study may have shown more significant results if the teachers and coach had continued to work together into a second year of paired outside-the-classroom professional development and inside-the-classroom coaching.

Another unexpected result was in the data surrounding the second research question concerned with use of ambitious practices. These results showed that coached teachers' composite IQA scores were not significantly better than comparison teachers' scores at the conclusion of the study. The test used to determine whether coached teachers' composite IQA scores were different from comparison teachers' scores was the Mann-Whitney nonparametric t-test. Like the Wilcoxan Signed-Rank test, it uses rank orders instead of a direct comparison of raw scores, but unlike the Wilcoxan, the Mann-Whitney does not attach signs to the rankings. For one group's scores to be deemed different from the other's, the sum of their ranks (1 through 10 in this case) has to be significantly less or significantly more than the sum of the ranks for the other group at the conclusion of the study. In this case, the sum of the ranks for the coached group was too close to the sum of the ranks for the comparison group and was not significant.

Regarding the unexpected result that coached teachers' composite IQA scores were not significantly better than comparison teachers' scores at the conclusion of the study, there were some unanticipated factors at work. Two of the comparison teachers, who work in the same school, began with two of the three highest composite IQA scores. After the coaching study, one of these two comparison teachers maintained the same composite IQA score, and one increased the composite IQA score by two points. This school building previously employed a mathematics and literacy coach. The individual who had been in the coaching role is now a Title I support teacher. However, during the outside-the-classroom follow-up PD sessions, it became apparent that the Title I support teacher was still acting as a coach for the two control teachers in her building (as well as some other teachers not attending PD). These two control teachers ended the study with the highest scores in the control group, and tied for second highest in the ranked order of scores used in the Mann-Whitney nonparametric t-test. There is no way to determine the impact of the Title I teacher on the teachers' composite IQA scores. There may have been no impact, meaning these two comparison teachers would have scored just as well on the IQA without the support they had at their school, but there may have been great impact, meaning that without the support, the teachers might not have scored as well on the IQA.

On the other hand, expected results for this study were seen regarding AR rubric scores. AR rubric scores were used as another measure of teachers' use of ambitious teaching practices. Coached teachers significantly improved their scores on the AR rubrics from before the coaching study to after the conclusion of the coaching study. Like the composite IQA scores, the Wilcoxan Signed-Rank test was used to analyze the AR rubric scores. However, there was more data for the AR rubric scores. The sample size was 20 because there were five coached teachers with four AR rubric scores each. Unlike the analysis of the composite IQA scores, it was not necessary for every one of the AR rubric scores to increase for the result to be significant. Eighteen of the 20 rubric scores either increased of remained the same, leading to a significant result for the change in AR rubric scores for the coached teacher group. Also as expected, coached teachers' AR rubric scores were significantly better than the comparison teachers' AR rubric scores at the conclusion of the study, as evidenced with a chi squared test of significance.

One possible reason that the results of the study varied depended on the grain size deals with the size and scope of the study itself. The study examined the practices of five coached teachers and five control group teachers, so it was quite small. The study entailed only four coaching cycles for each of the coached teachers spread over a time period of less than six months during one academic year, so its scope was also small. If the study had been expanded to include more teachers in the coaching aspect that was paired with the professional development, the results may have been more consistent between small-grained data and larger-grained data. If the study had been expanded over a longer time period, such as two or even three academic years instead of one, the results may have been more consistent (Campbell, 2012; Matsumura et al., 2013). If coaching had been more frequent, happening every few weeks as opposed to approximately once a month, perhaps differences from beginning to end or between coached and comparison group would have crystalized into consistently significant or consistently insignificant results.

Another possible reason for the inconsistency in results between larger and smaller grain sized data may involve an unmeasured variable. Teacher attitudes and beliefs were not measured at any point during the study. This may have impacted teachers' willingness to change instruction and consequently may have impacted whether instruction was changed and the degree to which instruction was changed to involve more ambitious teaching practices as measured via the IQA.

5.1.3 Context of the findings

With the results explained, this chapter now turns to contextualizing the results within some of the research reviewed in the literature review related to this study. In particular, this section contextualizes the coaching model, the results regarding changes in teacher use of ambitious practices, and the research design.

While the length of time for this dissertation's study was less than that of impactful coaching studies and consequently may have limited the results, a factor at play in this study that compares to other studies was the coaching model. The coaching model used for this study is closely aligned to many of the prevalent coaching studies in the literature. The coaching cycle used for this study involved a content-focused coaching model of co-planning, enacting, and debriefing. This coaching cycle is not only consistent with the content-focused coaching model developed at IFL and written about by West and Staub (2003), it also aligns with core portions of the coaching model used by Campbell and Malkus (Campbell & Malkus, 2011, 2014), the MIST study (Gibbons & Cobb, 2016), and the TN + IFL Mathematics Coaching Project (Russell et al., 2019). Also consistent with the larger TN + IFL Math Coach Model, the coach for this study had pedagogical and content goals in mind during coaching, and she assured that her coaching feedback was evidence-based. Like the Russell et al. project, this study also refined coaching as the study progressed. While the coach in this study did not specifically use of the Plan-Do-Study-Act cycle (Bryk et al., 2015), she did use teacher feedback and notes from coaching that provided insight into teachers' perceived needs to plan for the next coaching cycle. Thus, the coach in this study did adapt her coaching to needs–both needs expressed by the coachees and needs perceived by the coach–of the teachers with whom she was working.

There are also elements of the Tennessee + IFL model that are not in common with this study. The study under consideration in this dissertation did not examine the depth and specificity of coaching conversations, nor did this study examine the degree to which the coach maintained an inquiry stance with coachees. The coach in this study did not consistently think about the instructional triangle (D. K. Cohen et al., 2003) during coaching conversations, as is prescribed within the TN + IFL Math Coaching Model. However, the consistency of the coaching cycle used in this study with that of other studies in combination with the use of teacher expressed and perceived needs aligns this study with other impactful coaching studies.

Also aligned with some previously reviewed studies of professional development, were the results that coached teachers in this study did exhibit more ambitious teaching in mathematics, as measured via the AR rubric scores within the IQA toolkit. Research done in conjunction with the ESP program (Boston, 2013; Boston & Smith, 2009, 2011) found that teachers in the study changed their instruction by choosing and implementing more cognitively challenging tasks while better maintaining the level of demand. Teachers in the study for this dissertation also improved their selection of cognitively challenging tasks as measured on AR1: Potential of the Task and their sustenance of the cognitive demand, as measured on AR2: Implementation of the Task. These two rubrics from the IQA toolkit align with the effective Mathematics Teaching Practice of *implementing tasks that promote reasoning and sense making*.

As previously observed, while coached teachers in the study for this dissertation did significantly improve AR rubric scores, their composite IQA scores did not significantly improve. This differs from some previously reviewed studies of coaching. For example, Matsumura et al.'s (2013) results showed that coached teachers increased their mean IQA score when seven AR rubrics for literacy were averaged. Using an average of AR rubric scores as representative of an overall IQA score, as Matsumura et al. did, is commensurate with employing a composite IQA score that sums rubric scores, as this study did. Therefore, the Matsumura et al. study in literacy coaching showed more positive results compared to this study in mathematics coaching. Kraft, Blazar, and Hogan's (2018) meta-analysis found that pooled results for coaching studies had an overall positive effect size for classroom instruction. The study completed for this dissertation did not show an overall positive effect for classroom practices when the composite IQA scores are used as a measure.

Lastly, three elements of the research design of this study: pairing coaching with outside-the-classroom professional development; measuring the instructional practices of teachers; and using opportunities to learn about ambitious teaching practices as a dependent variable will be contextualized within current research. Firstly, this research study aligns with the design of some of the current research on coaching that uses or advocates for a coaching model pairing inside-the-classroom coaching with outside-the-classroom professional development. Within their five-part theory of action for improving mathematics teaching, Cobb and Jackson (2011) recommend pairing more formal teacher professional development with job-embedded professional development in the larger frame of a "coherent system of supports for ambitious instruction" (p. 9). Coaching is one part of the job-embedded PD recommended by Cobb and Jackson. Along with teacher networks, coaching is mentioned as a key component for "improving mathematics instruction at scale" (p. 9). Krupa and Confrey (2012) echoed this sentiment in their case study of a coach within their paired coaching and PD model. They said, "research has shown convincingly that teachers are not likely to change their instructional practices solely by attending isolated professional developments, and that ongoing support can help teachers implement the ideas presented in these professional developments" (p. 161). Neufeld and Roper (2003) also said coaching plus outside-the-classroom professional development could increase schools' instructional capacity, writing, "in light of our current knowledge about what it takes to change a complex practice like teaching, there are reasons to think that coaching, in combination with other professional development strategies, is a plausible way to increase schools' instructional capacity" (p. 1).

While recommendations for and discussion of pairing coaching with outside-the-classroom PD are within the cited research, none of these publications studied the changes in teachers' instructional practices when coaching was paired with outside PD as this study did. One consequential difference between this study and previous studies of coaching is the on-going nature of the outside-the-classroom PD for *both* the coaches and the comparison teachers. All teachers in the study attended ten days of summer PD, *and* all teachers in the study attended four follow-up sessions during the school year. Some studies, like Boston and Smith (2009, 2011), have shown that PD of this nature can make a difference in teachers' instructional practices. However, even with the on-going PD with its potential impact for all teachers, the coached teachers' AR rubric scores outpaced the scores of the comparison group in this study.

Another part of the research design for the study under consideration in this dissertation is aligned in part with that of Matsumura et al. (2013). The Matsumura group measured the instructional practice of teachers involved in coaching in their longitudinal study, as did the study for this dissertation. While Campbell (2012) did not directly measure teacher practice, she did measure teacher beliefs on a continuum from Traditional to Making Sense and used that as an indicator for teacher practice. However, many other recent coaching studies focused on the coaches and their coaching practices, not on the teachers and their teaching practices. For one, Campbell and Malkus (2014) followed teachers making the transition to coaching and documented changes in the coaches' content knowledge, MKT, and beliefs about mathematics teaching and learning. Gibbons and Cobb (2016) examined a case study coach's work in middle-school mathematics over a period of four years as she developed in her role of "support[ing] teachers' development of ambitious instructional practices" (p. 238). While the team isolated five important aspects of coaching practice and two important aspects of coaching knowledge, Gibbons and Cobb did not subsequently examine the teachers' practice in their study to determine if it became more ambitious. Russell[18] et al. (2019) discussed their coach training and model of coaching used in Tennessee in an effort to (a) analyze the research team and coaches' use of the continuous improvement model of adaptation and (b) determine which portions of the broad coaching model and more specific coaching framework resulted in "students' opportunities to engage in conceptual thinking" (p. 22).

The last element of the research design to be contextualized within the current research is the use of "opportunities to learn about ambitious teaching practices" as a variable. While there exist studies of coaching that aimed to measure the use of ambitious teaching (Matsumura et al., 2013), as this study does, there were no studies in the literature reviewed by this author that used teachers' opportunities to learn about ambitious teaching practices as a variable in the study. There were studies that measured the number of coaching sessions in which teachers partook, but that variable is different from measuring the number of exposures to ambitious teaching practices, as this study did. Related to the number of exposures to ambitious teaching practices in mathematics, this study also examined the depth of each exposure, and documented the specific teaching

---

[18] Russell et al. (2019) do not provide information about improved teacher practice in their publication. However, the coaching model for the TN + IFL Math Coaching Study does include "Improve Mathematics Teaching" as an output and lists four indicators (e.g., Engage students in productive struggle, Maintain cognitive demand of high level tasks) that align with four of the eight effective teaching practices for mathematics. One expects that future publications may address teacher practice.

practice(s) each learning opportunity exposed. Measuring opportunities to learn about ambitious teaching practices instead of the number of coaching sessions provided additional information about the content of the coaching. Knowing about the content of the coaching, specifically with regard to the effective Mathematics Teaching Practices addressed, provides the potential for better indicators of why or how teachers who changed their teacher practices did so. This leads to considering the implications of examining teachers' opportunities to learn about ambitious mathematics teaching practices as well as implications of other aspects of this study and its results.

5.1.4 Implications of the findings

The study forming the basis for this dissertation and the findings from this study have possible implications for the practice of coaching. This study implies that it is not just the act of coaching that helps teachers become more ambitious in their practice, it is the *purposeful* integration of coaching with outside-the-classroom professional development and with ideas teachers are currently considering or ready to consider regarding their students, in other words, the teacher's expressed and perceived needs.

Before commentary on the idea of purposeful integration of coaching with outside-the-classroom professional development and with where teachers are in the curriculum and in their learning progression about ambitious teaching, it is important to comment on three other aspects of this study that contribute to current research recommendations: (a) high-quality professional development, (b) pairing outside professional development with coaching, and (c) framing coaching within a set of practices associated with ambitious instruction. Firstly, as previously discussed, coached teachers in the study for this dissertation had many more opportunities to learn about the effective Mathematics Teaching Practices, and the opportunities to learn were qualitatively different from those of the comparison teachers. This implies that it was the content and quality of the coaching that made a difference for the coached teachers, not just participation in coaching. Concerning the value of the *content* in this experiment, the coaching that was provided in this study allowed teachers opportunities to learn about the eight effective teaching practices for mathematics, interact with them during the coaching cycle, and use them during instruction with the support of an instructional coach who was knowledgeable about those practices. Kennedy (1999) commented on this saying, “A program whose content is not valuable will not be improved by increasing the number of content hours, distributing contact hours over time, providing in-class visits, and so forth. Structural features alone provide no guarantee of improved teacher learning or of eventual benefit to students” (p. 6). This study reifies Kennedy’s statement.

Concerning *quality*, the coaching plus outside-the-classroom professional development provided during this study aligned with four of the five components of high-quality professional development as depicted by Desimone (2009). Those five components are (a) content focus; (b) active learning; (c) coherence; (d) duration; and (e) collective participation. The coaching for this study focused on pertinent mathematical and pedagogical content; had teachers as active and critical participants in the coaching cycle; was coherent from cycle to cycle as one debrief provided information that partially informed the next focus for coaching; was not a “one and done” occurrence. This study did not have the collective participation of all teachers in a school or even in a grade level at a school, so that quality of Desimone’s framework was not present.

The second item for commentary is the pairing of outside-the-classroom professional development with coaching. This study strengthened current recommendations regarding the pairing of outside professional development and coaching. As stated earlier in this chapter, Cobb and Jackson (2011) advocate for mathematics coaching that provides “job-embedded support for teachers' learning" (p. 9) as one of the five key components in their theory of action for improving mathematics teaching at scale along with "a coherent system of supports for ambitious instruction" (p. 9) that also includes district professional development. Neufeld and Roper (2003) say that the pairing can increase instructional capacity, and Krupa and Confrey (2012) echo that sentiment saying that professional development with on-going support can help with implementation of ideas from PD. This study demonstrated that teachers who were coached *and* attended outside-the-classroom PD significantly increased their AR rubric scores and had better AR rubric scores than uncoached teachers who only attended the outside-the-classroom PD. The idea of pairing the two forms of teacher learning is reinforced by the results of this study because the outside-the-classroom PD alone did not allow comparison teachers to make the same instructional changes towards ambitious mathematics instruction. The comparison teachers attended the on-going follow-up PD, like the coached teachers. However, teachers' participation in coaching *with* the outside PD was the catalyst that facilitated teachers' movement towards more ambitious practice.

The last aspect of this study corroborating current research recommendations is the framing of coaching within a set of practices associated with ambitious teaching. One can think of this set of practices in two ways: the set of *instructional* practices or the set of *coaching* practices. In this study, the set of *instructional* practices framing coaching was the eight effective Mathematics Teaching Practices. Again, this study lends to the strength of Cobb and Jackson's (2011) work. In their theory of action, Cobb and Jackson continuously mention that all five of the key components for improving mathematics instruction interconnect around a small set of specific "high-leverage instructional practices" (p. 16). This study did that with the eight effective Mathematics Teaching Practices and had some positive results. The study was purposeful about its use of these eight practices. The first coaching session for each coached teacher was framed around the effective teaching practice of *establishing learning goals to focus learning.* (See Figure 2.2 for a graphic representation of the eight effective Mathematics Teaching Practices.) Additionally, *every* coaching session involved the effective practice of implementing tasks that promote reasoning and problem solving, and most of these coaching sessions rated the as in-depth opportunities to learn about tasks. Thus, the practice of considering high-level tasks was ever-present during coaching. Then, the coaching in this study took on some individualization within the box in Figure 2.2 where *facilitating meaningful mathematical discourse* backgrounds four other effective mathematics teaching practices. As the four coaching sessions occurred, each teacher's needs were differentiated one from the other. Teachers' expressed and perceived needs were different, so coaching focused on different effective practices, as happens with improvement science (Bryk et al., 2015).

Considering a set of *coaching* practices, some researchers have recently written about effective coaching practices. For example, Gibbons and Cobb's (2016) case study identified five important coaching practices that included "(a) identifying long-term goals for teachers' development, (b) assessing teachers' current instructional practices, (c) locating teachers' current instructional practices on general trajectories of teachers' development, (d) identifying next steps for teachers' development, and (e) designing activities to support teachers' learning" (p. 246). Russell et al. (2019) also completed research pertaining to a small set of coaching practices. They found that having goals for teachers' pedagogical and content learning in mind during coaching and having deep and specific conversations with teachers were critical aspects of a coach's practice. This study substantiated both sets of findings. Coaching in this study considered learning goals for teachers as well as teacher's current knowledge and skill set when designing coaching. This connects to the first implication of this study that may add to the existing research base.

Kennedy (2016) commented on the idea of transfer, saying,

> PD programs typically meet with teachers outside of their classrooms to talk about teaching, yet they expect their words to alter teachers' behaviors inside the classroom. They are at risk for what Kennedy (1999) called the problem of enactment, a phenomenon in which teachers can learn and espouse one idea, yet continue enacting a different idea, out of habit, without even noticing the contradiction. (p. 3)

Pairing coaching with outside PD changes the scenario described by Kennedy, but it may not be just the paired nature of the coaching and outside-the-classroom professional development, nor is it just the quality of the professional development, it may be the *purposeful* integration of the outside and inside PD that makes the transfer of ambitious teacher practices happen more readily. During this study, portions of the four follow-up professional development sessions were sometimes purposefully and sometimes serendipitously coordinated with coaching cycles. For example, every coaching cycle involved selection and use of a cognitively challenging task. Three of the four follow-up professional development sessions also involved experiences with challenging tasks. Therefore, coached teachers were often concurrently thinking about and discussing high-level tasks in both the PD outside their classroom and during coaching. Perhaps as a related consequence of this, the AR2 rubric: Implementation of the task saw four of five coached teachers improve their score from before to after the study, while the last coached teacher maintained the score. The AR-Q rubric on the quality of teachers' questions also saw improvements in coached teachers' scores. Three of five coached teachers improved from low level scores to the highest score on this rubric. The two other coached teachers' scores remained the same with one of these teacher's scores being the highest possible score. Related to this outcome, the effective teaching practice of *posing purposeful questions* had received coordinated focus in the outside-the-classroom PD and coaching during the school year.

Adding further to this implication that purposeful integration made transfer occur more readily are examples of how coaching was "same but different" when compared to the outside-the-classroom PD. Coaching was the same because it focused on improving teaching practice, like the outside PD did, but coaching was different because of its direct applicability. So, purposefully integrating the two differing situations for learning may have been beneficial. For example, all coached teachers had a "same but different" encounter with the teaching practice around writing learning goals. While all teachers were introduced to learning goals in one of the follow-up PD sessions, the first coaching cycle's co-planning session also involved a discussion of the difference between learning goals and performance goals with examples from *Taking Action* (Huinker & Bill, 2017). Teachers had already written a goal for the lesson, but in all cases, it was a performance goal or a listing of content. The coach and teacher worked together, using Huinker and Bill's example, to craft an appropriate learning goal for the upcoming lesson. Thus, learning about the practice of *establishing goals to focus learning* was coordinated between outside-the-classroom PD and coaching. Perhaps this coordination made the practice more meaningful and readily transferred by coached teachers.

Other examples of a "same but different" encounters occurred with Coached teacher 3 around discourse, Coached teacher 1 around productive struggle, and Coached teacher 5 around representations. Even after a follow-up session concerning discourse, Coached teacher 3 was hesitant to try using small groups with her primary students. However, after the teacher and coach worked together to plan and demonstrate a turn-and-talk for the children during their math class, she moved towards using student groups. Without this "same but different" chance during coaching, to work through the details of getting students started with discourse, this teacher may not have begun to use this teaching practice in her classroom. Coached teacher 1 had an impactful encounter with productive struggle during one of her coaching cycles. Perhaps serendipitously, the coached teacher had just experienced a follow-up session involving the growth mindset (Dweck, 2006) when the coach and teacher began planning for the *Ford and Logan add 45+36* task (Illustrative Mathematics, 2016a). (See Appendix K.2.) During planning and teaching the task, the coached teacher had the chance to think about how to support students' productive struggle and then subsequently support that struggle in her class with the coach by her side. Coached teacher 5 had a "same but different" chance to learn about representations during a coaching cycle. She planned and implemented the *Rabbit Pens* task (Math Design Collaborative, 2015) with the coach's support. (See Appendix K.3.) The coach-teacher team decided to use popsicle sticks to allow children to model the pens. Representations were an important part of students' engagement in this task, so the effect of this encounter with the practice may have been different for this teacher than if she had only had the encounter with *using and connecting representations* from the follow-up PD. Even though both learning experiences involved discussion of multiple representations and the connections among them, without the coaching coordinated with the PD, the teacher may not have included concrete materials in the implementation of the *Rabbit Pens* task.

Grossman (2009) found that training for the profession of teaching lacks close approximations of practice. The in situ nature of coaching provides the close approximations, that is true, but it may be more than just the close approximation; it may be the connection between the close approximation of practice and the other learning concurrently taking place around the same concept in a different situation. This implication aligns with writing about the situative perspective by Putnam and Borko (2000) wherein they state that "cognition is (a) situated in particular physical and social contexts; (b) social in nature; and (c) distributed across the individual, other persons, and tools" (p. 4), and add that "the physical and social contexts in which an activity takes place are an integral part of the activity, and that the activity is an integral part of the learning that takes place within it" (p. 4).

There is another element of integration taking place here that may be just as important as the integration of outside-the-classroom professional development with coaching; it is the integration of coaching with where teachers are–meaning both where teachers are in their curriculum so that coaching is consistent with learning goals teachers have for students and where teachers are in terms of their perceived and expressed needs. For example, Coached teacher 2 realized he needed to work on questioning He noticed himself asking a lot of leading questions. He detected questions where he was "just leaving a place for the students to fill in the blank." He stated that he really needed to work on "asking more open questions, so students' thought processes would become more important and valued." During the implementation of the coaching lesson, he still asked multiple "fill-in-the-word" questions but commented on it during the post-coaching conversation, saying that during the lesson he caught himself and asked himself, "I wonder why I am asking so many of these low-level questions again?" So, he became consciously aware of this aspect of his teaching. This teacher was now ready to learn more about better questioning, and fortunately, the coaching was integrated with the teacher's need.

Another, perhaps more somber, implication of this study is the old adage that you "cannot get blood from a stone." If teachers are not ready to learn; if they are not willing to get onto a learning trajectory for ambitious teaching, then they will either not become ambitious in their practice or they will "play along" during coaching but not change their practice outside the coaching episodes. This was the case with Coached teacher 4. Coached teacher 4 collaborated with the coach during the coaching cycle. The teacher began to write learning goals for the coached lessons instead of performance goals. She accessed resources suggested by the coach to find cognitively challenging tasks for her students to engage in during the coaching sessions. While there was some struggle to do so, with the coach's guidance, the teacher did begin to ask questions without one-word answers that were pertinent to the lesson, and during coaching, students in the class engaged productively for a greater percentage of the class time. However, during the post-observation, there was little evidence of the progress made during coaching. Coached teacher 4 selected a low-level (procedures without connections) task for use with the students. The demand of the task did not increase upon implementation, as is usual with implementation of low-level tasks (Boston & Wilhelm, 2017). Any discussion that took place had students provide brief or one-word responses, and the questions asked by the teacher were formulaic in that the teacher asked the same question of all students instead of inquiring about their thinking or work. Thus, the conclusion is that those who continue not to want to make their teaching practice more ambitious, even after coaching, will not do so. Neufeld and Roper(2003a) comment on this aspect of coaching, saying, "They [coaches] can diagnose their learners' [the teachers'] needs and employ multiple coaching approaches; but, in the end, if the learner–either teacher or principal–does not or is not willing to learn, coaches cannot be successful" (p. 18).

On the other hand, for those teachers ready and willing to become more ambitious in their practice, coaching can make a difference. One of the coached teachers who made improvements had an interaction with his principal in the midst of his coaching experiences. The principal noticed that the teacher had made changes in his practice and that the teacher had become enthusiastic for taking on a leadership role around an upcoming district initiative regarding mathematics curriculum. The principal said to the coach, "This [the coaching and PD] has really been a turning point for you this year." The teacher replied to the principal that "This has been a turning point in my career." Thus, the anecdotal and qualitative information indicates that if teachers were ready and willing to work at becoming more ambitious in their teaching practice, they did. However, questions persist about whether teachers will continue becoming more ambitious in their instruction without coaching.

The final implication of this study deals with scalability. Scalability of a professional development model that integrates outside-the-classroom PD with in situ coaching is difficult to with a large number of teachers. Although the author believes this study provides an important way of studying the impact of coaching plus outside PD, the difficulty of providing more than a few coaching sessions per teacher must be acknowledged. If a scale-up of the model in this study were attempted, the researchers would need to consider trade-offs of what it would take to take such a model to scale.

5.2 Limitations

Certainly, scalability is one limitation that applies to this study, but the study for this dissertation has a few other limitations that likely affected the results. Firstly, the size of the study was indeed a limitation. Secondly, the time frame of the study limited its effectiveness. Lastly, use of only two groups of teachers: one coached and one uncoached, but both receiving outside-the-classroom professional development, may have limited the generalizability of the findings. The chapter addresses each of these limitations in turn.

5.2.1 Size of the study

This study involved a small number of teachers overall and small number of coached teachers. The study involved ten teachers total; five teachers were coached. With such a small sample, the generalizability of the findings is limited. Additionally, some statistical tests cannot be used or must be used in an altered form when sample size is small. For example, the Wilcoxan Signed-Rank test is altered for samples sizes of ten or fewer, and with a sample size of five coached teachers, all summed ranks must have consistent (positive or negative) signs for the test to show significant results. The variation of the chi-squared test for data sets where the expected valued may be less than or equal to five, called the Fisher Exact Probability test, had to be used for data analysis. Thus, the small size of his study served as a limitation for its utility.

5.2.2 Length of the study

This study reported on in this dissertation lasted less than one school year, and four coaching sessions occurred for each coached teachers The impactful studies of coaching (or of teacher support approximating a coaching model) mentioned in the literature review of this dissertation often lasted multiple school years and called for more frequent teacher-coach interactions. For example, the CGI study with elementary teachers used mentor teachers and CGI university staff who served as liaisons for the teachers in the study over three school years (Franke et al., 2001). The “type of support varied depending on the mentor and the teacher, but included observing in the teacher’s classroom and discussing the children’s thinking, planning lessons together, and assessing children together” (Franke et al., 1998, p. 21). The university liaison visited each teacher’s classrooms about every other week in the first two years of the study and monthly in the third year of the study. Over the life of the CGI study, teachers showed changes in classroom practices with “increased emphasis on problem solving, more communication by the children about their problem-solving strategies, and clear evidence that the teacher was more apt to attend to her own students' thinking” (Fennema et al., 1996, p. 415).

Campbell’s (2012) and Campbell and Malkus’ (Campbell & Malkus, 2011, 2014) studies also provide support for the idea that longer-term coaching has a greater impact. The study employed two different cohorts of coaches and one control group. One cohort of coaches worked with teachers for a three-year period. The other cohort of coaches worked with teachers for a one-year period before the conclusion of the study. The coaches who had been in their roles for three years had a greater impact on the teachers and the students of those teachers. Significant increases in student achievement for grades 3, 4, and 5 were not seen after the first year of a coach’s placement. The increase occurred as coaching became enculturated in the schools in years 2 and 3 of the coach’s placement (Campbell, 2012). “The pragmatic implication of this finding is the caution that a coach’s positive effect on student achievement develops over time” (Campbell & Malkus, 2011). The study done for this dissertation did not have the opportunity to allow the positive effect of the coach to develop over multiple years.

Campbell (2012) reported on changes in teacher beliefs as a stand-in for examining instructional changes in teachers’ practice, writing, “teachers’ perceptions of mathematics teaching and learning interrelate with their instructional practices…so change in teacher beliefs about mathematics teaching and learning is another way of evaluating the effect of specialists” (p. 155). The findings from this study show that “the beliefs of teachers who were highly engaged with a specialist changed significantly, shifting away from the Traditional perspective toward a Making Sense perspective” (p. 156), but the beliefs of teachers not “highly engaged” with a coach neglected to change in one direction or the other. While Campbell does not specify what "highly engaged" means with respect to frequency of coaching episodes, Matsumura, Garnier, and Spybrook (2013) do.

Like Campbell and Malkus (Campbell & Malkus, 2011, 2014), Matsumura et al. (2013) performed a longitudinal coaching study. Matsumura et al. also had two cohorts of coaches working with teachers. The first cohort of coaches was in place for three years, and the second cohort was in place for two years. The research team expected coaches to coach each of the teachers once a month and host grade-level meetings once a week. While most coaches did not fully meet this expectation, many teachers did participate in a level of activity close to the desired level. Teachers participated in coaching four to six times a year and met in grade levels at least once a month or more. While this study took place in literacy, not mathematics, its results had similarities to Campbell and Malkus'. By the end of the second year of the study, significant improvements occurred in teacher practice, as measured by the IQA, and significant improvements occurred in student achievement, as measured by the TAKS assessment. In comparison to the study for this dissertation, the more impactful studies occurred over multiple years with the accumulation of more coaching sessions or activities mirroring coaching.

Lastly, regarding the time period over which coaching occurs, Neufeld and Roper (2003) wrote "it will take several years for teachers to master what are fundamentally new and different instructional strategies even when teachers are eager to implement what they are learning" (p. 22). They wrote that the process will take even longer with reticent teachers. This leads to the possibility that the study reported on for this dissertation may have seen stronger, consistently significant results, if the time period of the study had been extended to multiple years.

5.2.3 Coaching capacity

Another limitation of this study was its restricted coaching capacity. One coach participated in this study. There was a total of ten teachers across the coached and comparison groups with eight schools involved. The impactful studies of coaching involved multiple coaches and larger numbers of teachers across numerous schools, and sometimes numerous districts. Russell et al. (2019) involved the state of Tennessee in its coaching initiative. The MIST study (Cobb & Jackson, 2011; Gibbons & Cobb, 2016) involved four large urban school districts. Campbell and Malkus' (Campbell, 2012; Campbell & Malkus, 2011, 2014) studies involved 36 schools. Matsumura et al. (2013) involved 29 schools. Having one coach in this study limited its capacity.

5.3 Recommendations for future research

Although this study adds to the overall picture of coaching's effectiveness as a form of teacher professional development, there is still much more than should be studied.

5.3.1 Design experiments

The TN + IFL Math Coaching Model (Russell et al., 2019) and the MIST project (Cobb & Jackson, 2011) both used a form of a continuous improvement model. Russell et al. used the Plan-Do-Study-Act (PDSA) cycle (Bryk et al., 2015). Cobb and Jackson's team used design experiments (Cobb, Confrey, DiSessa, Lehrer, & Schauble, 2003). One recommendation for future research in a small scale coaching experiment like the one described in this dissertation is to integrate more of the idea of the design experiments and the continuous improvement model. While Russell et al. used the PDSA cycle to refine their model of coaching, smaller scale studies can use the iterative nature of the PDSA cycle in conjunction with the coaching cycle of planning, enacting, and debriefing to purposefully allow the debrief to shape the planning and implementation of the next lesson. While the debriefing conversations in this study did shape the future coaching cycles, there was no extensive documentation recording the decision-making process in detail along with descriptions of rationales for the decisions made. Future studies can create and use tools or detailed notes as a means of being more explicit and purposeful about choices made for the next coaching cycle and reason for the choices made.

5.3.2 Development of effective teaching practices checklists

Another avenue for future research is the development of an observation tool for explicitly and specifically examining the eight effective teaching practices for mathematics based on the expected teacher behaviors as listed in *Principles to Action: Ensuring Mathematical Success for All* (NCTM, 2014). Although the IQA continues to be an effective tool for examining ambitious mathematics teaching and for examining elements of the eight effective Mathematics Teaching Practices, perhaps there is a place for a tool that parses the eight effective teaching practices for mathematics one from the other and is specific about measuring the occurrence of each. There exist IQA AR rubrics specifically measuring the teaching practices of *implementation of tasks*, *facilitation of discourse*, and *posing purposeful questions*. As a means of illustrating this direct connection, AR1:Potential of the Task is directly related to the effective Mathematics Teaching Practice called *implement tasks that promote reasoning and sense making*. One of the indicators for the effective Mathematics Teaching Practice is “posing tasks that require a high level of cognitive demand" (NCTM, 2014). To earn a 4 on AR1, the teacher must use a task rated as either a Doing Mathematics or Procedures with Connections (Boston, 2012c). These are the terms used in the Task Analysis Guide (Stein et al., 2000) to describe tasks that require a high level of cognitive demand. However, this sort of *direct* connection between the IQA and the effective Mathematics Teaching Practices is not ubiquitous.

Although the IQA measures the aforementioned practices and implicitly aligns with other effective Mathematics Teaching Practices, as evidenced in the rubric descriptors (e.g., *support productive struggle*), the IQA was developed prior to the publication of *Principles to Action* (NCTM, 2014). Thus, it could not have intentionally aligned with those exact indicators of ambitious mathematics teaching. In designing this study and performing the data analysis, the researcher perceived a need for a direct measurement of *all* of the effective Mathematics Teaching Practices. The researcher for this study attempted development of such a tool but did not complete the process. The Effective Teaching Practice (ETP) checklist (See Appendix I), was developed with some input from outside experts. The Effective Teaching Practices checklist directly relates to each effective mathematics teaching practice, but it was developed explicitly for this study. The tool was employed during the study solely by the researcher. No one substantiated the ratings teachers earned on these checklists for the pre-observation or post-observation. There was no opportunity for the repeated, iterative review and subsequent norming that goes into developing a valid and reliable research tool.

Getting to the point of qualitatively or quantitatively measuring as many as possible (or at least as many as possible that are not explicitly included in IQA rubrics), up to all eight, of the effective Mathematics Teaching Practices would take extensive additional work. Some of that work entails gathering group of mathematics educators to discuss the tool and its indicators; using the tool in mathematics classrooms with different coders in pilot tests; discussing ratings and reasons for ratings to attempt to reach consensus; reshaping indicators based on those conversations; and repeating the cycle until all could agree on indicators and on what a particular rating for that indicator looks like in the mathematics classroom. For example, in the current untested version of the ETP checklist, the indicators for productive struggle seem especially poorly developed and poorly articulated in the tool. In any future improvement cycle, these indicators would likely be the subject of extensive discourse.

Ultimately, this potential ETP checklist would have to answer the base question of "What in the tool indicates 'effective' or 'ambitious' for each of the eight practices and for the overall tool?" Determining what rating(s) or indicator(s) in the tool constitute effective mathematics instruction might also be arduous. One point for such a discussion may be whether to use the number of indicators exhibited, the average rating for the indicators, or both to make the determination of effectiveness or ambitiousness. One other potential point for discussion during the development of a tool such as the ETP checklist might be whether and how use the new tool in conjunction with the IQA to "drill down" on specific effective Mathematics Teaching Practices.

5.3.3 Additional comparison groups

All teachers involved in this study received the outside-the-classroom portion of the professional development experience. The study was designed to investigate how coaching paired with outside-the-classroom professional development facilitated a change in pedagogical practice. More specifically, the study measured differences between teachers who received coaching plus outside-the-classroom professional development and teachers who received only the outside-the-classroom professional development. There was no attempt to quantify or qualify any differences between teachers who received no professional development and the other groups of teachers. Nor was there any attempt to quantify or qualify differences between the groups and teachers who received *only* coaching without any other form of professional development.

Outside-the-classroom professional development may have been a factor in some of the control group teachers use of the effective Mathematics Teaching Practices as measured via the IQA AR rubrics. Perhaps there was some transfer of learning, contrary to previous findings for outside-the-classroom professional development provided by the MSC group (Removcik, 2014; Wang & Romero, 2013). The addition of other comparison groups in future studies: a coaching only group and a group with no professional development, may allow for findings that discriminate among more the factors responsible for the results.

5.3.4 More purposeful integration of out-of-class professional development with coaching

An area for continued study comes from the implication that purposeful coordination of coaching with outside-the-classroom professional development may have facilitated transfer of ambitious teaching practices to the classroom. If each of the effective Mathematics Teaching Practices received focus in outside-the-classroom professional development at nearly the same time as the practices received focus in coaching, might there be an even better opportunity for teachers to transfer and subsequently maintain the effective Mathematics Teaching Practice in question? Thus, a new question arises about whether being even more purposeful about the integration of the outside-the-classroom PD into the coaching cycles will have greater impact on teaching effectiveness.

Looking from the coaching end of the professional development experience, perhaps the outside-the-classroom PD could be altered to align better with coaching. An example of this might be purposefully using Case Stories (Hughes et al., 2008) in the professional development. Coached teachers could use one of their coaching episodes for their case story. The case stories from the participating teachers can be used as a determinant of which effective Mathematics Teaching Practice(s) become the foci of the follow-up session that day. In this way, there might be a more purposeful connection between the two forms of professional development.

5.3.5 Additional research

In chapter 1 of this dissertation, the author posed four possible explanations for why mathematics teachers do not adopt ambitious teaching practices. Those possible explanations were: (a) a lack of content knowledge for teaching; (b) use of low-level tasks or the lowering of task demands upon implementation; (c) attitudes and beliefs not commensurate with ambitious instruction; and (d) a lack of close approximations of practice (Grossman et al., 2009) during professional development. The study implicitly addresses the second reason concerning use and implementation of cognitively challenging tasks. Each of the coached teachers helped to find and implement high-level tasks for the coaching sessions in which they engaged, and their IQA AR1 and AR2 scores generally increased. However, the research questions for this study were not specifically about the use of tasks. The research questions concerned exposure to and use of ambitious mathematics teaching practices. This study explicitly addressed the fourth of these four possible explanations. Coaching provided the close approximations of practice via its in situ nature, and data was gathered concerning teachers' opportunities to learn about ambitious teaching practices. The argument might also be made that this study partially addressed the first issue surrounding a lack of MKT, but the data was not gathered to tackle that possible explanation for a lack of ambitious mathematics teaching.

Future research should more explicitly examine each of these four potential reasons for why teachers do not adopt ambitious teaching practices. Perhaps future studies might expand to gather data that addresses multiple reasons within one study. For example, this study might have employed questionnaires, or interviews, or a combination thereof, to measure teacher attitudes and beliefs. This study might have employed a post-post-assessment of MKT by re-administering the LMT assessment to the ten teachers in the study or to all of the teachers in the outside-the-classroom professional development.

Future research might also examine coaching practices in addition to teaching practices. The research might use the five key coaching practices espoused by Gibbons and Cobb (2016) to anchor research about coaching's effectiveness. Alternatively, research teams might examine the use of Russell et al.'s (2019) coaching framework and its effect on teacher practices in the classroom. Another option might be to pair a coaching framework – either Gibbons and Cobb's or Russell et al.'s – with the eight effective teaching practices for mathematics to track the effect of this pairing.

5.4 Concluding remarks

In an article confronting why teachers go back to the same pedagogical practices after participating in professional development as they had employed before participation in the professional development, Stein and Wang (1988) commented, "Over the past two decades, there has been continuing growth in the research base on what constitutes effective teaching" (p. 171). Three decades hence, the research community is still confronting the problem of why teachers do not consistently transfer learning from professional development to the classroom. The difference is that there now exists a set of well-defined practices that constitute effective teaching in mathematics. As Boston and Wilhelm (2017) wrote, “mathematics education research consistently identifies a set of instructional practices that appear to support students’ learning of mathematics with understanding, collectively called ‘ambitious mathematics instruction’ (Franke, Kazemi, & Battey, 2007)” (p. 830) This study employed that set of instructional practices for mathematics, now called the eight effective Mathematics Teaching Practices, in hopes of finding an avenue to catalyze a change in teachers’ pedagogical practice. The study sought to change teachers’ practice by pairing professional development outside teachers’ classrooms with content-focused coaching performed in the classroom.

The results from this study revealed a number of things regarding the pairing of outside professional development with coaching. Firstly, the coached teachers in this study had more opportunities to learn about effective teaching practices for mathematics than did their uncoached counterparts. Secondly, the coached teachers improved their scores on rubrics measuring ambitious mathematics teaching, and coached teachers scores’ were better than the comparison teachers’ scores at the conclusion of the study. Thirdly, the coached teachers’ experiences with the effective Mathematics Teaching Practices during coaching were qualitatively different than the experiences of the uncoached teachers, whose only opportunities to learn were during the outside-the-classroom professional development. These findings, with special focus on the third finding, connect to the an important implication of the study.

These findings not only imply that coaching matters, they also show that the content and the quality of what happens during coaching around a small set of well-defined ambitious practices (Cobb & Jackson, 2011) (e.g., the effective teaching practices for mathematics) makes a difference for teachers’ classroom practice. These findings also imply that pairing coaching with outside-the-classroom professional development focused on the same set of well-defined ambitious practices helps teachers better implement the teacher practices in their classrooms. This is especially true when the pairing *purposefully* integrates the foci of coaching with outside-the-classroom professional development insofar as ambitious teaching practices are concerned and purposefully integrates coaching with teachers' expressed and perceived needs. Coaching doesn't just need to happen; coaching needs to count.

Appendix A Comparison of rubrics and rubric labels used for instructional quality assessment toolkit for mathematics

Appendix A.1 Instructional quality assessment classroom observation rubric labels from three studies

Resnick et al. (2006)	Matsumura et al. (2008)	Boston (2012a)*
Academic Rigor (AR) rubrics	Cognitive Demand rubrics	Instructional Tasks/ Task Implementation
Potential of the task	Potential of the task	Potential of the task
Implementation of the task	Implementation of the task	Implementation
Student discussion of math concepts following task	Classroom Talk rubrics	Explanations of Mathematical Thinking and Reasoning
Accountable Talk (AT) rubrics	Rigor of discussion following the task	Rigor of the discussion
Student participation in the discussion	Student participation in the discussion	Participation
Teacher links student contributions to each other	Teacher links student contributions to each other	Teacher's linking
Students link to each other's contributions	Students link to each other's contributions	Students' linking
Teacher presses for evidence or for students to explain	Teacher presses for accurate knowledge and for students to explain	Teacher's press
Students give evidence or explain their thinking	Students provide accurate knowledge and explain their thinking	Students' providing
Clear Expectations (CE) rubrics	Teacher's Expectations rubrics	Teacher's Expectations

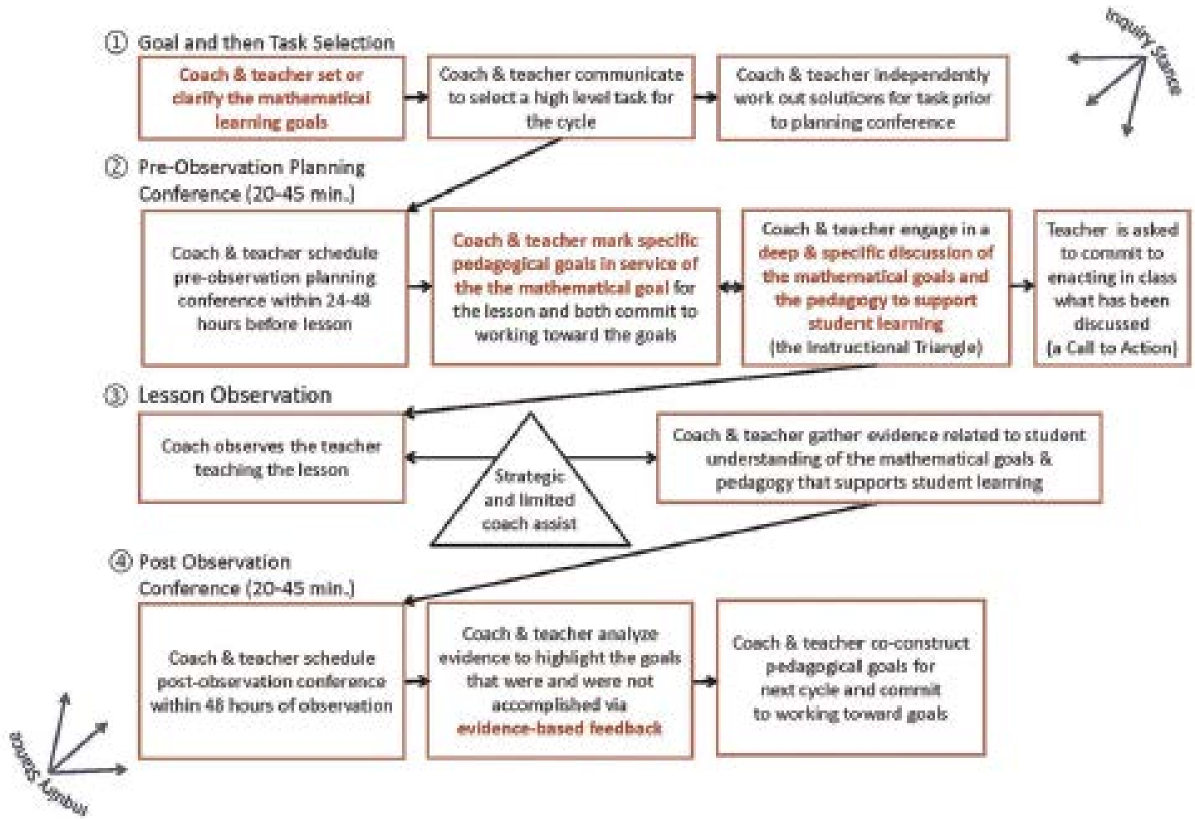
Clarity and detail of expectations	Clarity and detail of the expectations for student learning	Clarity and detail
Rigor of expectations	Rigor of the expectations for student learning	Rigor
Student access to expectations	Student access to expectations	Student access

***Note that while Boston's training tools (Boston, 2012d, 2012c, 2012b) use the AR, AT, and CE categories for the IQA rubrics, the article cited here (Boston, 2012a) used a different alignment to call out the IQA's alignment with four indicators of ambitious mathematics instruction.**

Appendix A.2 Instructional quality assessment assignments collections rubric labels from three studies

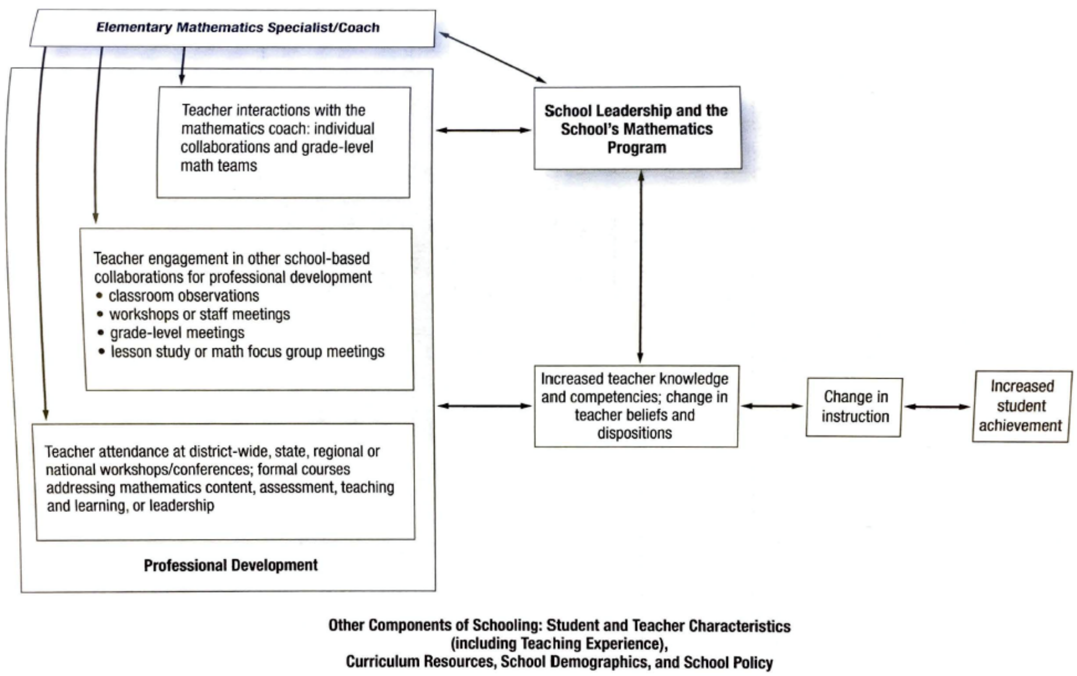
Resnick et al. (2006)	Matsumura et al. (2008)	Boston (2012a)
Academic Rigor (AR)	Cognitive Demand	Instructional Tasks/ Task Implementation
Potential of the task	Potential of the Task	Potential of the Task
Implementation of the task	Implementation of the Task	Implementation
Rigor in students' responses to the task	Rigor of student work following task	Explanations of Mathematical Thinking and Reasoning
Rigor in teacher's expectations		Rigor of students' written responses
Clear Expectations (CE)	Teacher's Expectations	Teacher's Expectations
	Rigor of the expectations for student learning	Rigor of teacher's expectations
Clarity and detail of expectations	Clarity and detail of the expectations for student learning	Clarity and detail
Student access to expectations	Student access to expectations	Student access

Appendix B Coach-teacher discussion process



Appendix Figure 1 Coach-Teacher Discussion Process (Russell et al., 2019, p. 7)

Appendix C Campbell's (2012) adaptation of Desimone's (2009) framework



Appendix Figure 2 How mathematics specialist/coaches influence professional development, classroom practice, and student learning (Campbell, 2012, p. 147)

Appendix D Sample items from learning mathematics for teaching

1. Ms. Dominguez was working with a new textbook and she noticed that it gave more attention to the number 0 than her old book. She came across a page that asked students to determine if a few statements about 0 were true or false. Intrigued, she showed them to her sister who is also a teacher, and asked her what she thought.

   Which statement(s) should the sisters select as being true? (Mark YES, NO, or I'M NOT SURE for each item below.)

	Yes	No	I'm not sure
a) 0 is an even number.	1	2	3
b) 0 is not really a number. It is a placeholder in writing big numbers.	1	2	3
c) The number 8 can be written as 008.	1	2	3

2. Imagine that you are working with your class on multiplying large numbers. Among your students' papers, you notice that some have displayed their work in the following ways:

```
Student A      Student B      Student C
    35             35             35
  × 25           × 25           × 25
  ----           ----           ----
   125            175             25
 +  75          + 700            150
  ----           ----            100
   875            875          + 600
                                 ----
                                 875
```

   Which of these students would you judge to be using a method that could be used to multiply any two whole numbers?

	Method would work for all whole numbers	Method would NOT work for all whole numbers	I'm not sure
a) Method A	1	2	3
b) Method B	1	2	3
c) Method C	1	2	3

3. Ms. Harris was working with her class on divisibility rules. She told her class that a number is divisible by 4 if and only if the last two digits of the number are divisible by 4. One of her students asked her why the rule for 4 worked. She asked the other students if they could come up with a reason, and several possible reasons were proposed. Which of the following statements comes closest to explaining the reason for the divisibility rule for 4? (Mark ONE answer.)

   a) Four is an even number, and odd numbers are not divisible by even numbers.
   b) The number 100 is divisible by 4 (and also 1000, 10,000, etc.).
   c) Every other even number is divisible by 4, for example, 24 and 28 but not 26.
   d) It only works when the sum of the last two digits is an even number.

(Hill & Ball, 2004, pp. 350-351)

Appendix E Excerpt from MSP-MSC survey: Beliefs and attitudes

Please indicate your opinion about each of the statements below:

	Strongly Disagree	Disagree	Neutral/Undecided	Agree	Strongly agree
Students learn mathematics best when they ask a lot of questions					
Students need to practice mathematical computation skills regularly to perform well on tests					
All students can learn challenging content in mathematics					
Students learn mathematics best in classes with students of similar abilities					
It is important for students to learn basic mathematics skills before solving problems					
I have adequate curriculum materials available for instruction					

Please indicate whether you agree or disagree with each statement.

	Strongly Disagree	Disagree	Uncertain	Agree	Strongly Agree
I am confident in my ability to use a variety of assessment techniques to evaluate students' mathematical learning and progress.					
Even when I try, I don't teach mathematics as well as I teach other subjects.					
I believe home environment has a greater effect on students' math achievement than my teaching.					
When a student has difficulty understanding a mathematics concept, I am usually at a loss as to how to help the student understand it better.					
I am confident in my ability to anticipate problems and confusions that students might have with particular math topics or concepts.					
I am confident in my ability to set appropriate math learning goals for the students in my class.					
	Strongly Disagree	Disagree	Uncertain	Agree	Strongly Agree
You can learn new things, but you can't really change your basic intelligence.					
I can assist families in helping their children understand or learn about mathematics.					
I understand mathematics concepts well enough to be effective in teaching mathematics.					
You can change even your basic level of talent considerably.					
I am confident in my ability to further students' math knowledge when they make spontaneous math comments or discoveries.					

(University of Pittsburgh (Collaborative for Evaluation and Assessment Capacity), 2016)

Appendix F Summary table of data for teachers in the study

Appendix Table 1 Summary data for teachers included in the study

Teacher	School type	Grade Level	LMT Score	Survey Score
Comparison Teacher 1	Urban-like	4	83	76
Comparison Teacher 2	Suburban	5	40	84
Comparison Teacher 3	Suburban	1	73	82
Comparison Teacher 4	Suburban	3	53	73
Comparison Teacher 5	Urban-like	1	50	80
Coached Teacher 1	Suburban	1	80	96
Coached Teacher 2	Rural	5	57	70
Coached Teacher 3	Suburban	1	57	82
Coached Teacher 4	Urban-like	2	73	82
Coached Teacher 5	Urban-like	3	37	81

Appendix G Opportunities to learn about effective teaching practices data collection tool

Opportunities to learn about effective teaching practices data collection tool

Name of Professional Development session (e.g. coaching #x, follow-up #y) Date:

Effective Teaching Practice (ETP)	Opportunity to Learn? (No; Some; or Extended)	Notes about opportunity to learn about ETPs
ETP 1: Goals	Example: SOME	Example: Coach created three possible goals statements related to the topic for the lesson, and together the coach and teacher analyzed each and chose the goal.
ETP 2: Tasks		
ETP 3: Representations		
ETP 4: Discourse		
ETP 5: Questions		
ETP 6: Procedural from Conceptual		

Effective Teaching Practice (ETP)	Opportunity to Learn? (No; Some; or Extended)	Notes about opportunity to learn about ETPs
ETP 7: Productive Struggle		
ETP 8: Evidence of Student Thinking		

Appendix H Academic rigor rubrics from the instructional quality assessment used for this study

Academic Rigor

RUBRIC 1: Potential of the Task

Did the task have potential to engage students in rigorous thinking about challenging content?

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. • students may be asked to engage in doing mathematics or procedures with connections, but the underlying mathematics in the task is not appropriate for the specific group of students (i.e., too easy <u>or</u> too hard to promote engagement with high-level cognitive demands); • students may need to identify patterns but are not pressed for generalizations or justification; • students may be asked to use multiple strategies or representations but the task does not explicitly prompt students to develop connections between them; • students may be asked to make conjectures but are not asked to provide mathematical evidence or explanations to support conclusions
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 requires no mathematical activity.**
N/A	Students did not engage in a task.

(Boston, 2012b, p. 9)

RUBRIC 2: Implementation of the Task
At what level did the teacher guide students to engage with the task in implementation?

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 is no explicit evidence of students' reasoning and understanding. • students identified patterns but did not form or justify generalizations; • students used multiple strategies or representations but connections between different strategies/representations were not explicitly evident; • students made conjectures but did not provide mathematical evidence or explanations to support conclusions • students engaged in doing mathematics or procedures with connections, but the underlying mathematics in the task was not appropriate for the specific group of students (i.e., too easy <u>or</u> too hard to sustain engagement with high-level cognitive demands).
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 connections to the concepts or meaning underlying the procedure being used.** Focus of the implementation 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	**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	**Students did not engage in mathematical activity.**
N/A	**The students did not engage in a task.**

(Boston, 2012b, p. 21)

RUBRIC 3: Student Discussion Following Task

To what extent did students show their work and explain their thinking about the important mathematical content?

4	Students show/describe written work for solving a task and/or engage in a discussion of the important mathematical ideas in the task. During the discussion, students: provide complete and thorough explanations of why their strategy, idea, or procedure is valid; students explain why their strategy works and/or is appropriate for the problem; students make connections to the underlying mathematical ideas (e.g., "I divided because we needed equal groups"). OR Students show/discuss more than one strategy or representation for solving the task, provide explanations of why/how the different strategies/representations were used to solve the task, *and/or make connections between strategies or representations. [Thorough presentation and discussion across strategies or representation]*
3	Students show/describe written work for solving a task and/or engage in a discussion of the important mathematical ideas in the task. During the discussion, students provide explanations of why their strategy, idea, or procedure is valid and/or students begin to make connections BUT the explanations and connections are not complete and thorough (e.g., student responses often require extended press from the teacher, are incomplete, lack precision, or fall short making explicit connections). OR Students show/discuss more than one strategy or representation for solving the task, and provide explanations of why/how the individual strategies/representations were used to solve the task *but do not make connections between different strategies or representations. [Thorough presentation and/or discussion of individual strategies or representations (but no cross-talk between different strategies or representations).]*
2	Students show/describe written work for solving the task (e.g., the steps for a multiplication problem, finding an average, or solving an equation; what they did first, second, etc) but do not engage in a discussion of why their strategies, procedures, or mathematical ideas work*; do not make connection to mathematical concepts. [Procedural explanations only]* OR Students show/discuss only one strategy or representation for solving the task.
1	Students provide brief or one-word answers (e.g., fill in blanks); OR **Student's responses are non-mathematical.**
0	There was no discussion of the task.
N/A	Reason:

(Boston, 2012b, p. 32)

Rubric AR-Q: Questioning	
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 2 times during the lesson, the teacher asks academically relevant questions (probing, generating discussion, exploring mathematical meanings and relationships).
2	There are one or more superficial, trivial, or formulaic efforts to ask academically relevant questions (e.g., 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) strong effort is made to ask academically relevant questions.
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:

- Intended to be used during whole group discussion.
- Alternatively, projects can decide to give a score for AR-q (small group work) and AR-Q (whole group work).

(Boston, 2012b, p. 39)

Appendix I Effective teaching practices checklists

Effective Teaching Practice 1: Goals

Establish mathematical goals to focus student learning: Effective teaching of mathematics establishes clear goals for the mathematics that students are learning, situates goals within learning progressions, and uses the goals to guide instructional decisions.	
Code	What are *teachers* doing?
	• Establishing clear goals that articulate the mathematics that students are learning as a result of instruction. To earn a +: ○ The established goal must be a *learning* goal and not simply a performance goal that articulates the enduring understanding students should gain or make progress towards, as a result of instruction. To each a 0: ○ The established goal can be a *performance* goal that states the skill or facts students are to gain/attain as a result of instruction. To earn a -: ○ There is no goal articulated for the lesson. Neither a learning goal nor a performance goal is established.
	• Using key ideas related to the mathematical learning goal during the enactment of the lesson. To earn a +: ○ The teacher is aware of and asks students about components of the mathematical learning goal or uses those key ideas or components to guide questions asked throughout the class. Examples may include asking assessing or advancing questions related to key ideas or important components of the learning goal (e.g. why certain figures that are closed are or are not classified as polygons; how adding 4+3 donuts can possibly give the same result as adding 3+4) or sequencing student solutions in a certain way to build towards understanding of the learning goal through connecting questions. To each a 0: ○ Although there is a learning goal established, the teacher ignores some or all of the underlying components or key ideas which are relevant to the class' task, small group work, or discussion. To earn a -: ○ The goal is a performance goal and not a learning goal. OR ○ There is no goal.
	• Using the mathematical learning goals to guide the class discussion. To earn a +: • Evidence exists that the teacher continued the planned instruction or adjusted the planned instruction in order to align instruction with the mathematical learning goal. To each a 0: • Although a learning goal has been established, the teacher seems to be using a performance or procedural goal to guide the discussion/discourse in the class. This may be evidenced in asking surface level questions or the same standard question or each student or group without taking their thinking into account. For example, asking students to only name the polygon as opposed to inquiring about its attributes OR

	• There is only a performance goal established, but the teacher *does* adjust instruction to students' progress (or the lack thereof) towards the performance goal. To earn a -: • Neither a learning goal nor a performance goal is established. OR • Evidence exists that the teacher ignored the goal (learning or performance) in order to stay in step with the curriculum guide or the book. OR • The activity or task was not aligned with the established goal.

Adapted from (NCTM, 2014, p. 16).

Coding: + → present in the form described in the bullet point(s)
0 → not fully engaged with the descriptor(s) for a (+) rating in this practice
- →detracts from the practice or from its intent

Effective Teaching Practice 2: Tasks

<table>
<tr><td colspan="2">Implement tasks that promote reasoning and problem solving: Effective teaching of mathematics engages students in solving and discussing tasks that promote mathematical reasoning and problem solving and allow multiple entry points and varied solution strategies.</td></tr>
<tr><td>Code</td><td>What are teachers doing?</td></tr>
<tr><td></td><td>• Selecting tasks that provide multiple entry points.

Point of clarification: Because this indicator uses the term “selecting,” we will consider only the task as it is written, not the task as it is implemented. This is akin to the IQA rubric on the Potential of the Task (Boston, 2012).

To earn a +:
o The task, as it is written (or as it is initially posed if it is not a written task), has more than one possible entry point including, but not limited to, multiple ways of representing the task (e.g. visual or symbolic) or multiple versions of the same type of representation (e.g. multiple symbolic representations can work). Thus, the task, as it is written, provides access for students.
o The task, as it is written, exhibits a level of cognitive demand that would earn a 3 or 4 on the Task Analysis Guide (Stein, Smith, Henningsen, & Silver, 2000) upon which the IQA rubric for Potential of the Task was based.

To each a 0:
o The task, as it is written (or as it is initially posed if it is not a written task), exhibits a level of cognitive demand that would earn a 2 on the Task Analysis Guide and on the IQA rubric for Potential of the Task.
o The task may elicit multiple versions of the same type of representation which might be used to solve, but the solutions themselves are limited to procedural strategies. For example, there may be more than one equation for solving a numeric problem, but any possible solution is purely procedural without connections to concepts or underlying meanings.

To earn a -:
o The task, as it is written, exhibits a level of cognitive demand that would earn a 1 on the Task Analysis Guide and on the IQA rubric for Potential of the Task OR there is no mathematical activity.
o There is only one desired entry point to answer the question at hand.</td></tr>
<tr><td></td><td>• Posing tasks that require a high level of cognitive demand.

Point of clarification: Because this indicator uses the term “posing,” we will consider the task as it is implemented. This is akin to the IQA rubric on the Implementation of the Task.

To earn a +:
o The task, as it is implemented, exhibits a level of cognitive demand that would earn a 3 or 4 on the IQA rubric for Implementation of the Task (Boston, 2012) because students are engaged in “exploring and trying to understand the nature of mathematical concepts, procedures and/or relationships” (p.6).
o The teacher works to maintain the level of complexity and cognitive demand that is in the original task as written when s/he launches the task with the class.

To each a 0:</td></tr>
</table>

	o The task, as it is implemented, exhibits a level of cognitive demand that would earn a 2 on the IQA rubric for Implementation of the Task because "(s)tudents are 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" (p. 6). o Teacher routinizes the task for the students, perhaps giving "hints" during the launch about how to solve the task. (e.g. "Make sure to make your candy jars look like the initial candy jar when you solve the task.") To earn a -: o The task, as it is implemented, exhibits a level of cognitive demand that would earn a 1 on the IQA rubric for Implementation of the Task because "(s)tudents are engaged in memorizing or reproducing rules, facts, formulae, or definitions" (p. 6). o Teacher does a sample problem of the same sort that students will do before the students have their individual or group time. o OR the students are not doing a mathematical activity/task.
	• Supporting students in exploring tasks without taking over student thinking. Point of clarification: Note that this is akin to the monitoring phase of the lesson as described in the *5 Practices for Orchestrating Effective Math Discussions* (Smith and Stein, 2018). Elements of the IQA rubric on the Rigor of Teachers' Questions (Boston, 2012) are relevant to this indicator as are the Factors of maintenance and decline (Stein et al, 2009). To earn a +: o The teacher circulates while students are working on the task. S/he asks academically relevant assessing and advancing questions of the students that "provide opportunities for students to elaborate and explain their mathematical work or thinking" (Boston, 2012, p. 8) in order that s/he might better understand students' current thinking and move them toward the learning goal. o Teacher scaffolding maintains or builds the complexity or cognitive demand of the task. o The teacher "sustain(s) press for justifications, explanations, and meaning through questions, comments, and feedback" (Stein et al, 2009, p. 16) To each a 0: (Not supporting the students in exploring, but also not taking over the thinking) o The teacher circulates during the task implementation. S/he asks students questions, however, these questions are usually procedural, factual, or superficial in nature. o There are few, if any, instances of questions that aim to understand students' thinking about the task at hand. o Students may be told to explain, but unclear explanations are accepted and/or procedural explanations are accepted and/or even desired. o Students may be held accountable for products or processes but they are not high level in nature. For example, a student is "wrong" because they did not use the correct plural for vertex. o Teacher scaffolding, when it is provided, may detract from the complexity or cognitive demand of the task. To earn a -: (Not supporting students in exploring, AND taking over the thinking.) o The teacher may not circulate during the task implementation, instead waiting for students to finish before going over correct answers, or what has been deemed the correct solution strategy for the problem(s). o The teacher may circulate and correct students' papers as s/he goes, possibly providing what s/he has deemed as the proper or correct way to complete the problem.

	○ The teacher may perform what he or she deems as the difficult portions of the task for students; thus, taking over the student thinking. ○ There may not be a true monitoring phase of the lesson because the teacher is providing step by step instructions to students about how to do the problem, occasionally stopping for students to fill in the blank s/he provides with the correct one word or short phrase answer.
	• Encouraging students to use varied approaches and strategies to make sense of and solve tasks. Point of clarification: Note that some of these descriptors are also related to or drawn from the Factors of maintenance and decline. To earn a +: ○ Teacher helps students draw conceptual connections to encourage a variety of ways of making sense of or thinking about the problem. To each a 0: ○ While students are "allowed" to use any approach they want to use to solve the posed problem, either ▪ it is clear (implicitly or explicitly) that the teacher favors one strategy over the other possibilities (e.g. "Remember that we want to use strategies that will give us the answer quickly.") OR ▪ no conceptual connections are drawn, even when different strategies are shared. To earn a -: ○ The teacher is dictating which strategy students should use during the task/problem set. ○ Teacher discourages students from using any approach aside from the one s/he has dictated. ○ Strategies other than the strategy espoused by the teacher or book but that may still apply to the problem at hand, are dismissed. Students may be told, "That is not what we are working on today."

Adapted from (NCTM, 2014, p. 22)

Coding: + → present in the form described in the bullet point(s) 0 → not fully engaged with the descriptor(s) for a (+) rating in this practice - →detracts from the practice or from its intent

Effective Teaching Practice 3: Representations

Code	What are *teachers* doing?
	Use and connect mathematical representations. Effective teaching of mathematics engages students in making connections among mathematical representations to deepen understanding of mathematics concepts and procedures and as tools for problem solving. Note that "representations refer both to process and to product – in other words, to the act of capturing a mathematical concept or relationship in some form and to the form itself….Moreover, the term applies to processes and product that are observable externally as well as to those that occur 'internally' in the minds of people doing mathematics" (NCTM, 2000, p. 67).
	• Selecting tasks that allow students to decide which representations to use in making sense of the problems. To earn a +: ○ The chosen task allows for multiple different representations (i.e., visual, symbolic, verbal, contextual, physical) to be used or multiple forms of the same representation e.g., three different visual representations) that are accessible to students. The teacher allows for or encourages the use of any or many of the representations. ○ Students decide for themselves which of the relevant representations makes the most sense for the problem at hand. To each a 0: ○ While the chosen task, itself, might allow for multiple representations, the teacher does not allow students to decide which to use. Instead s/he explicitly tells students which representation to use in solving the problem. (e.g., "We will use a number line to solve the problem today.") ○ The chosen task is written in such a way that one particular representation is called for or is an obvious implicit choice (e.g., a ten-frame appears next to the problem). OR ○ The teacher implies which representation should be used (e.g. based on recent instruction). To earn a -: ○ The chosen task or the teacher explicitly dictates which representation should be used and only one representation is deemed appropriate.
	• Allocating instructional time for students to use, discuss, and make connections among representations or within a single representation. To earn a +: ○ Adequate time and opportunity is provided for students to employ, talk about, and connect representations both during the monitoring/exploration and during the debrief/discussion. ○ Students may be explicitly asked about connections or there may be an established classroom routine or protocol for students to discuss and connect representations following the whole-class debrief. To each a 0: ○ Either students are given adequate time for exploration during the monitoring phase and a cursory or surface-level discussion takes place OR ○ Some time is provided for use and discussion and then for connection, but the time provided is inadequate in both phases of the lesson.

	To earn a -: ○ There is no instructional time provide for students to use, discuss, or make connections among representations. If there are any representations seen in the lesson, they are provided by the teacher.
	• Focusing students' attention on the structure or essential features of mathematical ideas that appear, regardless of the representation. To earn a +: ○ Students are asked about the relationship between the underlying structure or important features of the mathematical idea that their representation elicits or clarifies and the representation that they are using. To each a 0: ○ Students may be asked to explain their representation, but there is no questioning about how the representation relates to any mathematical ideas or learning goals. To earn a -: ○ Students are either not asked about their representation OR ○ Their representation is dismissed.

Adapted from (NCTM, 2014, p. 29)

Coding: + → present in the form described in the bullet point(s)
0 → not fully engaged with the descriptor(s) for a (+) rating in this practice
- →detracts from the practice or from its intent

Effective Teaching Practice 4: Discourse

<table>
<tr><td colspan="2">Facilitate meaningful mathematical discourse. Effective teaching of mathematics facilitates discourse among students to build shared understanding of mathematical ideas by analyzing and comparing student approaches and arguments.

Note: "The discourse in the mathematics classrooms gives students opportunities to share ideas and clarify understanding, construct convincing arguments regarding why and how things work, develop a language for expressing mathematical ideas, and learn to see things from other perspectives" (NCTM, 2014, p. 29).</td></tr>
<tr><td>Code</td><td>What are teachers doing?</td></tr>
<tr><td></td><td>
Engaging students in purposeful sharing of mathematical ideas by selecting and sequencing varied student approaches and solution strategies for whole-class analysis and discussion.
To earn a +:
The teacher purposefully selects varied student approaches to share during the whole-class debrief and sequences them in a way that can potentially further student progress towards the leaning goal.
Students share the strategies, and a mathematically relevant discussion occurs.
To each a 0:
Student solution strategies are shared, and may even be shared in a pre-selected and sequenced way, but no discussion is engendered by the sharing, so no analysis or discussion of the strategies takes place. OR
Students do share solution strategies in some random order with only the teacher's voice providing any analysis, instead of the student voices
To earn a -:
No student solution strategies are shared. Any student sharing is that of answers, next steps in the procedure, or are in the IRE pattern of questioning. OR
Student solutions beyond that described in the previous bullet are shared only by the teacher.
</td></tr>
<tr><td></td><td>
Ensuring progress toward mathematical goals by making explicit connections to the key mathematical ideas in the lesson.
To earn a +:
Students are encouraged, via questions or other means, to connect their thinking about the day's task to the mathematical learning goal (or key mathematical ideas, if no learning goal is provided).
To each a 0:
Only the teacher connects the solution strategies to each other or to the important math of the day. There is no evidence that students have made the connections between the day's work and the important math ideas therein.
To earn a -:
There are no connections made to the key mathematical ideas, even if discussion does occur.
</td></tr>
</table>

Adapted from (NCTM, 2014, p. 35)

Coding: + → present in the form described in the bullet point(s) 0 → not fully engaged with the descriptor(s) for a (+) rating in this practice - →detracts from the practice or from its intent

Effective Teaching Practice 5: Questions

<table>
<tr><td colspan="2">Pose purposeful questions. Effective teaching of mathematics uses purposeful questions to assess and advance students' reasoning and sense making about important mathematical ideas and relationships.

Note: "…questions that encourage students to explain and reflect on their thinking … help students make important mathematical connections, and support students in posing their own questions" (NCTM, 2014, p. 35-36).</td></tr>
<tr><td>Code</td><td>What are teachers doing?</td></tr>
<tr><td></td><td>• Advancing student understanding by asking questions that build on, but do not take over or funnel, student thinking.
To earn a +:
o Teacher questions advance the student thinking by using evidence of what the student already understands (because the teacher has elicited such information).
o Questions are broad but focused on the important mathematics in order not to imply or prescribe a particular path. (e.g., "The number sentence you wrote for today's scenario is 45 + 36 = 711. Why or how does 711 make sense as the sum of 45 and 36? …Tell me more about what the number 711 means in terms of tens and ones.")
To each a 0:
o Although the teacher asks probing questions that might lead students to a correct response, they are effectively taking over the student thinking by providing too much scaffolding along the way to an explanation. The teacher is funneling the student thinking. (e.g., How would you say this number: 711….How many 10s do you have? How much is 7 tens? What if you had 7 tens and 3 ones, how much is that? How much is 7 tens and 11 ones?)
o Teacher interacts with students, does not ask questions, but also does not tell the students about the math they are to learn. (e.g., S/he restated the activity directions.)
To earn a -:
o Instead of asking questions, teachers tell students about the math they are to learn.</td></tr>
<tr><td></td><td>• Making certain to ask questions that go beyond gathering information to probing thinking and requiring explanation and justification.
To earn a +:
o Students are explicitly asked for an explanation or a justification for their response or for their current thinking about the problem.
o Teacher presses the students to communicate their thoughts clearly.
To each a 0: (Simply gathering of information.)
o Any questions that are asked are able to be answered with short, one-word answers like "Yes", "No", "perpendicular", or "42."
o The teacher asks only questions where students are to fill in the missing word in his or her sentence. (e.g., "When we combine things, we have to ____." Or "When we see a number in this form, it is called ___.")
o Teacher questioning follows the Initiate-Response-Evaluate (I-R-E) pattern.
Note: Teachers may ask a mixture of questions that do go beyond information gathering and ones that are short-answer. There is no "cut off" number for how many questions that go beyond gathering information must be asked. Judge whether the rating is (+) or (0) what the predominant feel of the lesson is. What has more weight in the lesson?
To earn a -:
o No questions are asked of the students. The teacher tells the information that students need. OR Any questions that are asked are answered by the teacher in a rhetorical manner. (e.g., What does 711 tell you? Well, it tells you there are 7 hundreds, 1 ten, and 1 one. Is</td></tr>
</table>

	that what you get when you add 45 and 36? Of course, not. You can't get 700 from adding 40 and 30.)
	• Asking intentional questions that make the mathematics more visible and accessible for student examination and discussion. To earn a +: o Students are asked about the mathematical structures that can help them make connections and see mathematical relationships. (e.g. What does your array tell us about the problem situation for the band concert?) o Teacher's questions have the potential to help students build their level of sophisticated thought about the mathematics. (e.g., Instead of counting on by ones to add 24 and 16, what other tools might help you add 16 to 24?) o Students are asked to reflect on or connect their thinking to others in order to bring better or deeper understanding of the mathematical ideas. (e.g., How is your array related to Jim's equation?) To each a 0: o Teachers may be asking questions about student thinking, but the questions asked do not help make the mathematics more clear, visible, or accessible. To earn a -: o No questions are asked. o Any questions asked do not deal with the mathematics. o Questions actually obscure the mathematics of the lesson.
	• Allowing sufficient wait time so that more students can formulate and offer responses. To earn a +: o Sufficient wait time (at least 3 seconds) is provided for students both after a question is asked and after the initial response is provided to engender more and deeper student responses. To each a 0: o Some wait time is provided, but it is insufficient for students' needs. o Wait time may be initially provided when a question is asked, but once the initial response is garnered, either ▪ no other student responses are allowed or elicited, ▪ the teacher takes the thinking back from the students, OR ▪ other responses are gathered in rapid succession, so no deeper understanding is engendered. To earn a -: o No wait time is given or teacher fill in the silence with the response they desire instead of allowing for student responses.

Adapted from (NCTM, 2014, p. 41)

Coding: + → present in the form described in the bullet point(s)
0 → not fully engaged with the descriptor(s) for a (+) rating in this practice
- →detracts from the practice or from its intent

Effective Teaching Practice 6: Fluency

<table>
<tr><td colspan="2">Build procedural fluency from conceptual understanding. Effective teaching of mathematics builds fluency with procedures on a foundation of conceptual understanding so that students, over time, become skillful in using procedures flexibly as they solve contextual and mathematical problems.

Note: This practice promotes "an integrated and balanced development of concepts and procedures" (NCTM, 2014, p. 42) akin to that promoted by NMAP (2008) and NRC (2001). Thus, there are two indicators geared primarily towards the development of conceptual understanding and two indicators geared primarily towards the use of procedures.</td></tr>
<tr><td>Code</td><td>What are teachers doing?</td></tr>
<tr><td></td><td>• Providing students with opportunities to use their own reasoning strategies and methods for solving problems.[1]

To earn a +:

o Students are given the chance to choose among methods or strategies that make sense to them in order to solve mathematical problems. The methods may be self-created strategies or ones that classmates or the teacher have previously shared.

To each a 0:

o Students can choose a problem solving method, but only from a list that has been pre-determined by the teacher or the text.
o Students can choose a problem solving method, but only for one portion of the task and not for other portions of the task.

To earn a -:

o Students cannot choose their own strategy or method for problem solving.</td></tr>
<tr><td></td><td>• Using visual models to support students' understanding of general methods. (In part, this likely is based on research from Fuson and Beckmann (2012/2013) in which they say that "Standard algorithms are to be understood and explained and related to visual models before there is an focus on fluency" (p. 28).)

To earn a +:

o Teacher elicits the use of a visualization of the student's method via the use of manipulative models, other tools, or drawings/pictures in order to support students in understanding their method or that of a classmate.
o Students are asked about how their method is shown in the visual model.

To each a 0:

o Teacher provides a visual model of the student's method to him/her.

To earn a -:

o No use of visual model to help with understanding of general methods.</td></tr>
</table>

[1] Some lessons will focus more on the development of the foundation of conceptual understanding. For those lessons, the first two indicators are generally more applicable.

	• Asking students to discuss and explain why the procedures that they are using works to solve particular problems.[2]
	
	To earn a +:
	
	o Students are asked about
	  ▪ why the procedure used is appropriate and/or productive for the problem at hand (e.g., Why does this work?) or
	  ▪ about what the procedure is accomplishing for the student or
	  ▪ about what kind of results to expect as a result of using the procedure. (Martin, 2009).
	
	To each a 0:
	
	o Students are asked for a step-by-step run-through of the procedure.
	
	To earn a -:
	
	o Students asked only about the answer that resulted from performing the procedure or algorithm.

	• Connecting student-generated strategies and methods to more efficient procedures as appropriate. To earn a +: o Students asked about a "more efficient" or "quicker" way to perform the operation or procedure in question. (e.g., I see you shared 72 cookies among 6 kids one by one to divide 72 by 6. That seemed to work for you, but it took a long time. I wonder how else you might share the cookies with each kid to make it take a little less time.) o Students asked to compare or contrast their method to that of another in order to help develop efficiency. (e.g., I noticed when you divided 72 by 6, you created 6 circles and wrote a 10 in each and then put an x in each for another one. Compare your method to my imaginary friend, Jan's way of thinking about this problem. How is yours the same as his? How does your show things differently than his?) To each a 0: o Student-generated strategies are shared, but there is no connection to more efficient approaches, even though the lesson is focused on the use of procedures to solve problems and not on concept building. o Student-generated strategies are shared, but the teacher (or another student) then shows what has been pre-determined as the way the class should use. To earn a -: o Student-generated strategies are not allowed to be used. Only the teacher or book-determined "correct" algorithm or procedure is allowed.

Adapted from (NCTM, 2014, p. 47-48)

Coding: + → present in the form described in the bullet point(s)
0 → not fully engaged with the descriptor(s) for a (+) rating in this practice
- →detracts from the practice or from its intent

[2] Other lesson will focus more on the use of procedures to solve problems. For those lessons, the second two idicators are more applicable.

Effective Teaching Practice 7: Productive Struggle

Support productive struggle in learning mathematics. Effective teaching of mathematics consistently provides students, individually and collectively, with opportunities and supports to engage in productive struggle as they grapple with mathematical ideas and relationships.	
Code	What are *teachers* doing?
	• Giving students time to struggle with tasks. To earn a +: o Students are provided with time to mull over the task, to make potential missteps and realize they are unproductive, to try different representations or tools, to discuss the validity of their ideas, etc. To each a 0: o Teacher expects students to come up with a viable strategy within a few minutes of the problem being posed. To earn a -: o There is no time given to struggle; either the task is of such low cognitive demand that students immediately find an answer or the teacher leads the students through any potential struggle that might occur before students get a chance to think about the task.
	• Asking questions that scaffold students' thinking without stepping in to do the work for them. To earn a +: o Teacher questions keep the thinking and reasoning at a high level of cognitive demand. o Teachers ask students to explain and justify their reasoning and/or solution strategies, pressing students as needed to get quality explanations/justifications. To each a 0: o Teacher asks questions of the students, but the teacher's questions provide too much scaffold and/or effectively take over the thinking and remove the struggle. To earn a -: o Any questions are fact-based. o The teacher "rescues" the student by "breaking down the task and guiding students step by step through the difficulties" (NCTM, 2014, p. 48). o Teacher tells instead of asking questions. o Teacher only rereads or restates directions.
	• Beginning student interactions by inquiring about students' thought processes, and checking back with students to determine if progress was made after the initial interaction. To earn a +: o Teacher begins the questioning with an assessing question in order to inquire about the student's thought process without implying whether the student is correct or incorrect. o Teacher progresses to asking an advancing question and provides time for students to think before s/he circles back around to determine what progress student has made in the intervening time. To each a 0: o Teacher inquires about students thought process but then either

	• Does not ask an advancing question to keep them moving towards the learning goal OR • Never manages to circle back to the student(s) to check on any progress that was (not) made. To earn a -: o Teacher never inquires about students' thought processes; instead the teacher either checks only the answers, does not circulate while students are working, or circulates only as a classroom management strategy.
	• Praising students for their efforts in making sense of mathematical ideas and/or perseverance in reasoning through problems. (i.e. Providing useful feedback to students.) To earn a +: o Teacher provides specific and descriptive feedback to students related to their struggle with a particular task. (e.g., I see that you have tried using a number line to solve this problem, but you seem to be struggling with what to count by. Keep going with your thinking, and make sure to share your ideas with your partner(s).) To each a 0: o Teacher feedback is only neutral in supporting struggle with no feedback related to their particular struggle. (e.g., Keep working. I see the smoke coming from your ears.) o Teacher acknowledges the struggle after it has concluded. To earn a -: o No teacher feedback is provided OR o Feedback is praising the right answer or the intelligence of the student. (e.g., Good job. You are right.; You are so smart. I knew you would get it.)

Adapted from (NCTM, 2014, p. 52)

Coding: + → present in the form described in the bullet point(s)
0 → not fully engaged with the descriptor(s) for a (+) rating in this practice
- →detracts from the practice or from its intent

Effective Teaching Practice 8: Student Thinking

<table>
<tr><td colspan="2">Elicit and use evidence of student thinking. Effective teaching of mathematics uses evidence of student thinking to assess progress toward mathematical understanding and to adjust instruction continually in ways that support and extend learning.</td></tr>
<tr><th>Code</th><th>What are teachers doing?</th></tr>
<tr><td></td><td>
Eliciting and gathering evidence of student understanding throughout the lesson.
Note: Teachers will need to think ahead about what that evidence might look like with regard to the task and learning goals. As Principles to Action states, “Preparation of each lesson needs to include intentional and systematic plans to elicit evidence that will provide ‘a constant stream of information about how student learning is evolving toward the desired goal’” (NCTM, 2014, p. 53).
To earn a +:
Teacher chooses or structures a task or asks questions to “draw out specific understanding, conceptual gaps, or common errors” (p. 54) throughout the lesson.
To each a 0:
Teacher gathers evidence of student understanding
at the conclusion of the lesson only (e.g., via exit cards) OR
in a robotic or repetitive way that is non-specific to the task or the student.
To earn a -:
Teacher does not gather evidence of student understanding throughout or at the conclusion of the lesson.
</td></tr>
<tr><td></td><td>
Making in-the-moment decisions on how to respond to students with questions and prompts that probe, scaffold, and extend.
To earn a +:
Teacher responds to students’ understandings, or the lack thereof, with questions and/or prompts that support their current thinking and/or press it forward (e.g., ask students to restate the problem in their own words, change the numbers in the problem to easier or more difficult ones, compare and contrast solution methods)
To each a 0:
Teacher responds to students’ understandings or the lack thereof by taking over the thinking for the student and/or instructing him/her about what to do next.
To earn a -:
Teacher does not respond to student understandings or the lack thereof, instead, carrying on with the lesson.
</td></tr>
<tr><td></td><td>
Using what you have learned about students’ thinking via teacher-student interactions to move students forward in their thinking.
To earn a +:
Providing advancing questions and/or feedback to move students further towards the learning goal.
To each a 0:
Teacher does interact with students with respect to the task, but fails to take any action to move students forward in their thinking. (e.g. Teacher praises students or encourages them to engage with the task.)
</td></tr>
</table>

	○ Teacher does interact with students with respect to the task, but fails to take any action to move students forward in their thinking. (e.g. Teacher praises students or encourages them to engage with the task.) ○ Teacher's interactions with the students are repeating the same instructions or directions. (e.g., Make sure to start at 0 on your ruler when you measure.) To earn a -: ○ Teacher does not interact with students regarding their thinking and therefore, cannot move their thinking forward towards the learning goal.

Adapted from (NCTM, 2014, p. 56)

Coding: + → present in the form described in the bullet point(s)
0 → not fully engaged with the descriptor(s) for a (+) rating in this practice
- →detracts from the practice or from its intent

Appendix J List of effective teaching practices encountered in follow-ups

Appendix Table 2 Depth of encounters with effective Mathematics Teaching Practices in follow-up PD

	Depth of opportunity to learn			
Effective Mathematics Teaching Practice	Follow-up 1	Follow-up 2	Follow-up 3	Follow-up 4
Goals				some
Tasks		extended	some	some
Representations	some	some		
Discourse	extended	extended	some	some
Questions	some	some	some	some
Procedural from Conceptual				
Productive Struggle			extended	
Evidence of Student Thinking	some			extended

Appendix K Sampling of tasks used during coaching and outside-the-classroom professional development

Appendix K.1 Maria's money

Maria's Money

a. Ali had $9. Maria had $5. How many more dollars did Ali have than Maria?
Ali had $9. Maria had $5. How many fewer dollars did Maria have than Ali?

b. Ali had $4 more than Maria. Maria had $5. How many dollars did Ali have?
Maria had $4 less than Ali. Maria had $5. How many dollars did Ali have?

c. Ali had $4 more than Maria. Ali had $9. How many dollars did Maria have?
Maria had $4 less than Ali. Ali had $9. How many dollars did Maria have?

http:illustrativemathematics.org

Illustrative Mathematics

5.NF Scaling Up and Down

Task

The fifth grade teachers are in charge of planning the annual Davis Elementary Fun Run. The teachers decide that each adult should run $\frac{6}{4}$ as far as each student in grade 5 and each student in grade 1 should run $\frac{3}{4}$ as far as each student in grade 5.

a. Who has to run the longest distance? Who has to run the shortest distance? Explain your reasoning.

b. The fifth grade students decide that they should each run four laps around the track. How many laps should each adult and each first grade student run?

c. Peyton, a fifth grader calculates that he will run a $\frac{1}{2}$ mile. Write two multiplication equations involving $\frac{1}{2}$, one that shows how many miles each adult will run and one that shows how many miles each first grade student will run.

d. When Peyton showed the adults his calculations, some of them were confused. Some of the adults thought multiplication always makes a number larger, for example 2×5 is bigger than 5. When calculating the distance the first graders ran, Peyton used multiplication but got a smaller number. Explain why the product of 5 and another number is not always greater than 5, and write an example to help the adults understand.

e. Presley, another fifth grade student, wanted to write the distance she ran in eighths. She noticed that you could write this equation: $\frac{4}{4} \times \frac{1}{2}$miles = $\frac{4}{8}$ miles. Explain why in this case multiplying by $\frac{1}{2}$ results in a product that is neither larger nor smaller than $\frac{1}{2}$.

Appendix K.3 Ford and Logan add 45 + 36

Illustrative Mathematics

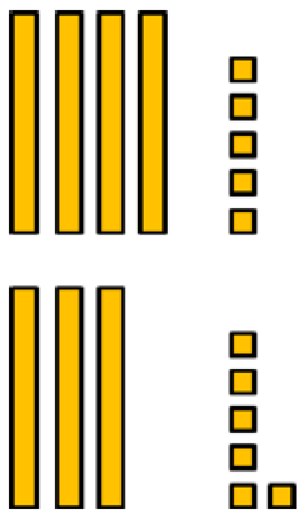
1.NBT Ford and Logan Add 45+36

Task

Actions

Part One: Solve the problem and explain your thinking.

45+36



Part Two: Ford and Logan each solved the problem using a different strategy.

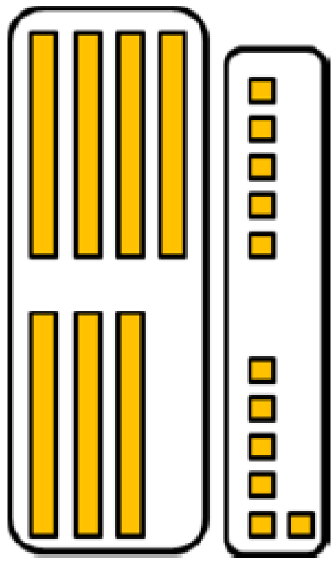
a. How did Ford solve the problem? Will his strategy always work?

b. How did Logan solve the problem? Will her strategy always work?

c. How are their strategies similar or different?

d. How was your strategy similar or different than Ford or Logan's?

Ford's thinking:



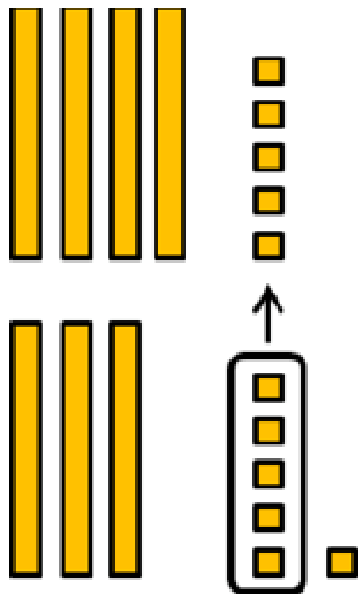
I Counted the tens first, so 10, 20, 30, 40, 50, 60, 70.

Then I counted the ones, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81.

So 45+36=81.

Logan's thinking:

Illustrative Mathematics



First I broke 36 into 30+1+5.

Then I gave 5 from 36 to the 45 to make 50 because 50 is a friendly number.

Then I added 30+50 to make 80. Then I added 1 to 80 to get 81.

So 45+36=81.

Appendix K.4 Building a rabbit pen task

Name ____

Building Rabbit Pens

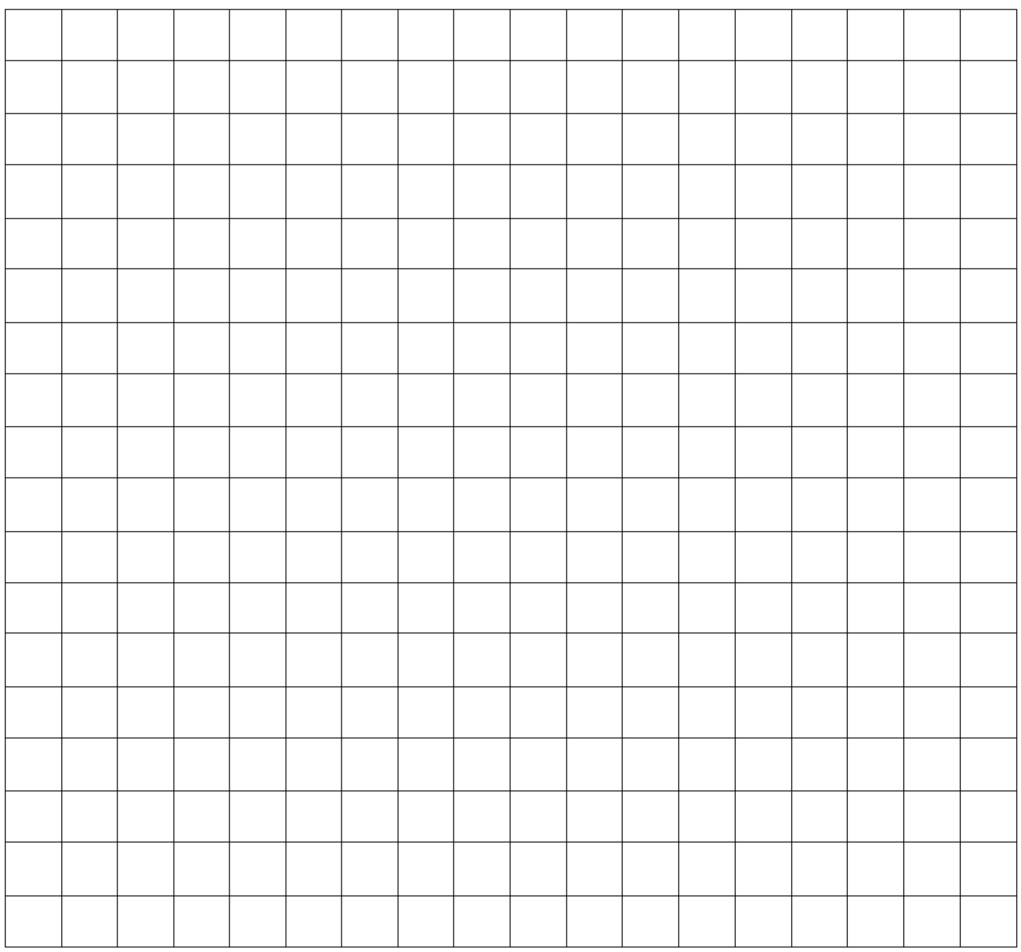
Antoine wants to build a rectangular pen for his rabbits.

He has 24 feet of fence that he can use to make his pen.

He plans to use all 24 feet of fence to make the best pen he can for his rabbits.

Use the grid to create some possible rabbit pens that Michael could build, making sure to label the pens.

Antoine wants his rabbits to have lots of space to run around. Which of the pens should Antoine build?



On the back of this sheet, explain WHY you think the pen you chose is the best one for Antoine's rabbit.

Appendix K.5 Joey's run

Joey's Run

Maya runs 1/2 of a mile each day for 12 days. How many miles does Maya run in all?

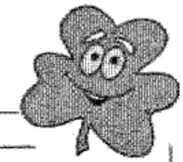
K.C. runs 1/4 of a mile each day for 12 days. How many miles does K.C. run in all?

Joey runs 3/4 of a mile each day for 12 days. How many miles does Joey run in all?

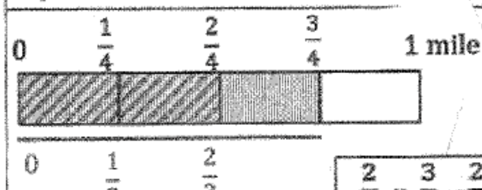
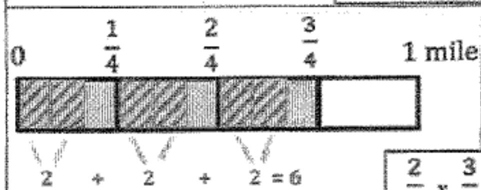
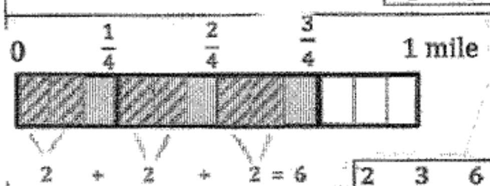
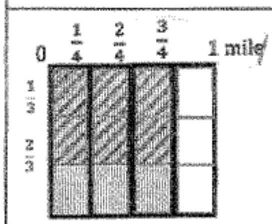
(tasks.illustrativemathematics.org)

Appendix K.6 Shamrock smile mile

Shamrock Smile Mile **Name:**



Put checks beside any of the representations that are accurate for our problem 2/3 of ¾. Select any 2 of the representations and write an explanation for why it is accurate or why it is not an accurate representation.

Representations	Explanation
 $\frac{2}{3} \times \frac{3}{4} = \frac{2}{4} = \frac{1}{2}$ of a mile	
 $\frac{2}{3} \times \frac{3}{4} = \frac{6}{9}$ of a mile	
 $\frac{2}{3} \times \frac{3}{4} = \frac{6}{12} = \frac{1}{2}$ of a mile	
 $\frac{2}{3} \times \frac{3}{4} = \frac{6}{12}$ $\frac{6}{12} = \frac{1}{2}$ mile	

(www.teachingchannel.org)

Appendix K.7 Box of clay

Illustrative Mathematics

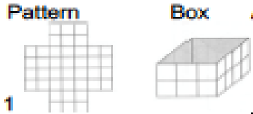
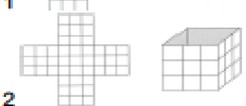
5.MD Box of Clay

Task

A box 2 centimeters high, 3 centimeters wide, and 5 centimeters long can hold 40 grams of clay. A second box has twice the height, three times the width, and the same length as the first box. How many grams of clay can it hold?

Appendix L Benchmark tasks grid

Benchmark Tasks Grid – Elementary

Look at the two addition strategies below. See if you can figure out how they work. Louisa's Strategy / Li's strategy 37 + 44 = ? / 37 + 44 = ? 37 + 40 = 77 / 40 + 44 = 84 77 + 4 = 81 / 84 – 3 = 81 37 + 44 = 81 / 37 + 44 = 81 **Now try to use either Louisa's Strategy or Li's Strategy to solve these problems:** 29 + 56 = ? 65 + 27 = ? **Which strategy do you think is easier? Explain.**	**Solve the following problems using any material that will help you find the answer.** Find the dimensions of a box that will hold twice as many cubes as a box that is 2 by 6 by 4. 1. Volume of original box:________________ 2. Volume of new box:________________ 3. Dimensions of new box:________________ **Explain how you found the dimensions of the new box.**	**Find as many different ways as you can to make this equation true.** 40 x 32 = _____________ x ____________ **How are the new numbers in your new expressions related to 40 x 32?**
Meredith uses an interesting strategy for solving subtraction problems when you have to trade. Try to figure out how it works. 42 – 27 / 34 – 19 On my first step, I get 12. / On my first step, I get 14. On my 2nd step I get 15. / On my 2nd step I get 15. 15 is my final answer. / 15 is my final answer. 71 – 36 First step:_____________ Second step:_____ Final answer:__________	**How many cubes fit in each box? First, determine the number of cubes without building the box. Then build a box and use cubes to check. Check your first answer with your actual answer before going on to the next box.** Pattern / Box / FIRST Answer / Actual 1. Box 1  ________ ________ 2. Box 2  ________ ________	**Write multiplication equations, solve the problems, and show your solutions.** • There are 4 people sitting at my table. Each person has 5 fingers on each hand. How many fingers are there altogether? • There are 12 people in my group. Each person has 2 eyes. How many eyes re there altogether? • Write a story problem that represents 4 x 3.
Solve each problem. Show your work. Write an equation. 1. Franco and Sally have 18 cherries and 13 grapes. How many pieces of fruit do they have?	Find the volume of each figure. 1.  2. 	Find each product. 12. 17 × 6 13. 28 × 6 14. 39 × 3 15. 16 × 8
Which number combination does not make 20? a. 10 + 10 b. 18 + 2 c. 19 + 3 d. 5 + 5 + 5 + 5	**Choose the best word to complete each sentence.** 1. The number of cubic units in a solid figure is the . 2. The point where 3 edges of a solid figure meet is a___. 3. The number of square units in a region is the ______. **Vocabulary** area vertex volume perimeter	**Multiply each number in the first column of the table with the number at the top. Circle any combinations you do not know immediately.** **Table A** (× 7): 2 – 14; 6 –; 8 –; 3 – **Table B** (× 8): 2; 9; 4; 11 **Table C** (× 6): 10; 4; 2; 8 **Table D** (× 9): 5; 2; 12; 4

Bibliography

Askey, R. (2001). Good intentions are not enough. In T. Loveless (Ed.), *The great curriculum debate: How should we teach reading and math?* (pp. 660–664). Washington, D.C.: Brookings Institution Press. https://doi.org/10.1086/603733

Ball, D. L., & Bass, H. (2000). Interweaving content and pedagogy in teaching and learning to teach: Knowing and using mathematics. *Multiple Perspectives on the Teaching and Learning of Mathematics*, (2000), 83–104.

Ball, D. L., & Forzani, F. M. (2009). The Work of Teaching and the Challenge for Teacher Education. *Journal of Teacher Education*, *60*(5), 497–511. https://doi.org/10.1177/0022487109348479

Ball, D. L., Hill, H. C., & Bass, H. (2005). Knowing mathematics for teaching: Who knows mathematics well enough t o teach third grade. *American Educator*, (Fall), 14–17, 20–22, 43–46. Retrieved from www.aft.org/pubs-

Ball, D. L., Thames, M. H., & Phelps, G. C. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, *59*(5), 389–407. https://doi.org/10.1177/0022487108324554

Bell, C. A., Wilson, S. M., Higgins, T., & Mccoach, D. B. (2010). Measuring the Effects of Professional Development on Teacher Knowledge: The Case of Developing Mathematical Ideas. *Journal for Research in Mathematics Education*, *41*(5), 479–512. https://doi.org/10.2307/41110411

Blazar, D., & Kraft, M. A. (2015). Exploring mechanisms of effective teacher coaching: A tale of two cohorts from a randomized experiment. *Educational Evaluation and Policy Analysis*, *37*(4), 542–566. https://doi.org/10.3102/0162373715579487

Boaler, J. (2006a). How a detracked mathematics approach promoted respect, responsibility, and high achievement. *Theory Into Practice*, *45*(1), 40–46. https://doi.org/10.1207/s15430421tip4501_6

Boaler, J. (2006b). Urban success: A multidimensional mathematics approach with equitable outcomes. *Phi Delta Kappan*, *87*(5), 364–369.

Boaler, J. (2013). Ability and Mathematics: the mindset revolution that is reshaping education. *FORUM*, *55*(1), 143–152. Retrieved from www.wwwords.co.uk/FORUM

Boaler, J. (2015). Jo's TEDx Stanford Talk. Retrieved November 7, 2018, from https://www.youcubed.org/resources/jos-tedx-talk/

Boaler, J., & Staples, M. (2008). Creating mathematical futures through an equitable teaching approach: The case of Railside school. *Teachers College Record*, *110*(3), 608–645. https://doi.org/Article

Borasi, R., & Fonzi, J. (2002). *Professional development that supports school mathematics reform*. *Foundations: A monograph for professionals in science, mathematics, and technology education* (Vol. 3). Arlington, VA: National Science Foundation. Retrieved from http://hdl.handle.net/2027/pst.000054512061

Borko, H., Kieran, C., & Lester, F. K. (2004). Teaching and learning. In F. K. Lester & J. Ferrini-Mundy (Eds.), *Proceedings of the NCTM research catalyst conference* (pp. 163–169). Reston, VA: National Council of Teachers of Mathematics.

Boston, M. (2012a). Assessing Instructional Quality in Mathematics. *The Elementary School Journal*, *113*(1), 76–104.

Boston, M. (2012b). Instructional Quality Assessment Classroom Observation Tool.

Boston, M. (2012c). Instructional Quality Assessment Classroom Observation Tool Accountable Talk Function Reference List, 1–13.

Boston, M. (2012d). IQA Mathematics Rubrics: Academic Rigor Training Manual.

Boston, M. (2013). Connecting changes in secondary mathematics teachers' knowledge to their experiences in a professional development workshop. *Journal of Mathematics Teacher Education*, *16*(1), 7–31. https://doi.org/10.1007/s10857-012-9211-6

Boston, M., Bostic, J., Lesseig, K., & Sherman, M. (2015). A Comparison of Commonly Used Mathematics Classroom Observation Protocols. *Mathematics Teacher Educator*, *3*(2), 154–170. https://doi.org/10.5951/mathteaceduc.3.2.0154

Boston, M., Dillon, F., Smith, M. S., & Miller, S. (2017). *Taking action: Implementing effective mathematics teaching practices in grades 9-12*. (M. Smith, Ed.). Reston, VA: National Council of Teachers of Mathematics.

Boston, M., & 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. https://doi.org/10.2307/40539329

Boston, M., & Smith, M. S. (2011). A "task-centric approach" to professional development: Enhancing and sustaining mathematics teachers' ability to implement cognitively challenging mathematical tasks. *ZDM - International Journal on Mathematics Education*, *43*(6), 965–977. https://doi.org/10.1007/s11858-011-0353-2

Boston, M., & Wilhelm, A. G. (2017). Middle school mathematics instruction in instructionally focused urban districts. *Urban Education*, 1–33. https://doi.org/10.1177/0042085915574528

Bransford, J., Brown, A. L., Cocking, R. R., & National Research Council. (1999). *How people* *learn: Brain, mind, experience, and school*. Washington, D.C: National Academy Press.

Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. G. (2015). *Learning to Improve: How America's Schools can get Better at Getting Better*. Cambridge, MA: Harvard Education Press.

Campbell, P. F. (2012). Coaching and elementary mathematics specialists. In J. M. Bay-Williams & W. R. Speer (Eds.), *Professional Collaborations in Mathematics Teaching and Learning: Seeking Success for All, Seventy-fourth Yearbook* (pp. 147–161). Reston, VA: National Council of Teachers of Mathematics.

Campbell, P. F., & Malkus, N. N. (2011). The impact of elementary mathematics coaches on student achievement. *The Elementary School Journal*, *111*(3), 430–454. https://doi.org/10.1086/657654

Campbell, P. F., & Malkus, N. N. (2014). The mathematical knowledge and beliefs of elementary mathematics specialist-coaches. *Zdm*, *46*(2), 213–225. https://doi.org/10.1007/s11858-013-0559-6

Carpenter, T. P., & Fennema, E. (1991). Research and cognitively guided instruction. In E. Fennema, T. P. Carpenter, & S. J. Lamon (Eds.), *Integrating research on teaching and learning mathematics* (pp. 1–16). Albany, NY: State University of New York Press.

Carpenter, T. P., Fennema, E., Franke, M. L., Levi, L., & Empson, S. B. (1999). *Children's Mathematics: Cognitively Guided Instruction*. Portsmouth, NH: Heinemann.

Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C.-P., & Loef, M. (1989). Using Knowledge of Children ' s Mathematics Thinking in Classroom Teaching : An Experimental Study. *American Educational Research Journal*, *26*(4), 499–531.

Chapin, S. H., O'Connor, C., & Anderson, N. C. (2009). *Classroom Discussions: Using Math Talk to Help Students Learn* (Second Edi). Sausalito, CA: Math Solutions.

Chval, K. B., Arbaugh, F., Lannin, J. K., van Garderen, D., Cummings, L., Estapa, A. T., & Huey, M. E. (2010). The transition from experienced teacher to mathematics coach: Establishing a new identity. *The Elementary School Journal*, *111*(1), 191–216. https://doi.org/10.1086/653475

Cobb, P., Confrey, J., DiSessa, A. A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, *32*(1), 9–13.

Cobb, P., & Jackson, K. (2011). Towards an empirically grounded theory of action for improving the quality of mathematics teaching at scale. *Mathematics Teacher Education and Development*, *13*(1), 6–33.

Cohen, D. (1990). A revolution in one classroom: The case of Mrs. Oublier. *Educational Evaluation and Policy Analysis*, *12*(3), 311–329. https://doi.org/10.3102/01623737012003311

Cohen, D. K., & Ball, D. L. (1999). *Instruction, capacity, and improvement*. Philadelphia: University of Pennsylvania.

Cohen, D. K., & Ball, D. L. (2001). Making change: Instruction and its improvement. *Phi Delta Kappan*, *83*(1), 73–77. https://doi.org/10.1177/003172170108300115

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

Confer, C. (2006). Being a successful math coach: Ten guiding principles. In C. Felux & P. Snowdy (Eds.), *The Math Coach Field Guide: Charting Your Course*. Sausalito, CA: Math Solutions.

Darling-Hammond, L., & Bransford, J. (Eds.). (2005). *Preparing Teachers for a Changing World: What Teachers Should Learn and Be Able to Do*. San Francisco, CA: Jossey-Bass.

Desimone, L. (2009). Improving Impact Studies of Teachers' Professional Development: Toward Better Conceptualizations and Measures. *Educational Researcher*, *38*(3), 181–199. https://doi.org/10.3102/0013189X08331140

Dweck, C. S. (2006). *Mindset : the new psychology of success* (1st ed.). New York, NY: Random House.

Dweck, C. S. (2007). The Secret to Raising Smart Kids. *Scientific American Mind*, (November), 1–6. Retrieved from http://users.digitalkingdom.org/~rlpowell/public_media/raising_smart_kids.pdf

Fennema, E., Carpenter, T. P., Franke, M. L., Levi, L., Jacobs, V. R., & Empson, S. B. (1996). A longitudinal study of learning to use children ' s thinking in mathematics instruction. *Journal for Research in Mathematics Education*, *27*(4), 403–434.

Findell, B. (2002). *Adding it up: Helping children learn mathematics*. *Book Reviews*. Retrieved from ftp://129.132.148.131/EMIS/journals/ZDM/zdm026r1.pdf

Fosnot, C. T. (2007). *Contexts for learning mathematics: The big dinner, multiplication with the ratio table*. Portsmouth, NH: Heinemann.

Foster, D., & Noyce M.D., P. (2004). The Mathematics Assessment Collaborative: Performance testing to improve instruction. *Phi Delta Kappan*, (January), 367–374.

Fraivillig, J. L., Murphy, L. A., & Fuson, K. C. (1999). Advancing Children's Mathematical Thinking in Everyday Mathematics Classrooms. *Journal for Research in Mathematics Education*, *30*(2), 148. https://doi.org/10.2307/749608

Franke, M. L., Carpenter, T. P., Fennema, E., Ansell, E., & Behrend, J. (1998). Understanding teachers' self-sustaining, generative change in the context of professional development. *Teaching and Teacher Education*, *14*(1), 67–80. https://doi.org/10.1016/S0742-051X(97)00061-9

Franke, M. L., Carpenter, T. P., Levi, L., & Fennema, E. (2001). Capturing teachers' generative change: A follow-up study of professional development in mathematics. *American Educational Research Journal*, *38*(3), 653–689. Retrieved from http://www.jstor.org.pitt.idm.oclc.org/stable/3202495

Fuson, K. C., Carroll, W. M., & Drueck, J. V. (2000). Achievement Results for Second and Third Graders Using the Standards-Based Curriculum Everyday Mathematics. *Journal for Research in Mathematics Education, 31*(3), 277–295. https://doi.org/10.2307/749808

Fuson, K. C., Smith, S. T., & Cicero, A. M. Lo. (1997). Supporting Latino First Graders' Ten-Structured Thinking in Urban Classrooms. *Journal for Research in Mathematics Education*, *28*(6), 738. https://doi.org/10.2307/749640

Fuson, K. C., Wearne, D., Hiebert, J., Murray, H. G., Human, P. G., Olivier, A. I., … Carpenter, T. P. (1997). Children's Conceptual Structures for Multidigit Numbers and Methods of Multidigit Addition and Subtraction. *Journal for Research in Mathematics Education*, *28*(2), 130. https://doi.org/10.2307/749759

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Suk Yoon, K. (2001). What makes professional development effective ? Results from a national sample or teachers. *American Educational Research Journal*, *38*(4), 915–945. https://doi.org/10.3102/00028312038004915

Gibbons, L. K., & Cobb, P. (2016). Content-focused coaching: Five key practices. *The Elementary School Journal*, *117*(2), 237–260.

Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P. W. (2009). Teaching Practice: A Cross-Professional Perspective. *Teachers College Record*, *111*(9), 2055–2100. https://doi.org/10.1177/0022487109336543

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. https://doi.org/10.2307/749690

Hiebert, J., & Carpenter, T. P. (1992). Learning and teaching with understanding. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 65–97). New York, NY: MacMillen.

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

Hiebert, J., & Grouws, D. A. (2007). The Effects of classroom mathematics teaching on students' learning. In F. K. Lester, Jr. (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 371–404). Charlotte, NC: Information Age Publishing.

Hiebert, J., Morris, A. K., Berk, D., & Jansen, A. (2007). Preparing Teachers to Learn from

Teaching. *Journal of Teacher Education*, *58*(1), 47–61. https://doi.org/10.1177/0022487106295726

Hiebert, J., Stigler, J. W., Jacobs, J. K., Givvin, K. B., Garnier, H. E., Smith, M. S., & Gallimore, R. (2005). Mathematics teaching in the united states today (and tomorrow): Results from the TIMSS 1999 video study. *Educational Evaluation and Policy Analysis*, *27*(2), 111–132. https://doi.org/doi:10.3102/01623737027002111

High-Leverage Practices. (n.d.). Retrieved from http://www.teachingworks.org/work-of-teaching/high-leverage-practices

Hill, H. C., & Ball, D. L. (2004). Learning mathematics for teaching : Results from California's Mathematics Professional Development Institute. *Journal for Research in Mathematics Education*, *35*(5), 330–351. Retrieved from http://www.jstor.org/stable/30034819

Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Content Knowledge : Conceptualizing and measuring teachers ' topic-specific knowledge of students. *Journal for Research in Mathematics Education*, *39*(4), 372–400. https://doi.org/Article

Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical Knowledge for Teaching and the Mathematical Quality of Instruction: An Exploratory Study. *Cognition and Instruction*, *26*(4), 430–511. https://doi.org/10.1080/07370000802177235

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of Teachers ' Mathematical Knowledge for Teaching on Student Achievement. *American Educational Research Journal*, *42*(2), 371–406. Retrieved from http://www.jstor.org/stable/30034819

Hill, H. C., Shilling, S. G., & Ball, D. L. (2004). Developing Measures of Teachers Mathematical Knowledge for Teaching. *The Elementary School Journal*, *105*(1), 11–30.

Horizon Research Inc. (2013). *2012 National survey of science and mathematics education: Highlights Report*. Chapel Hill, NC.

Hughes, E. K., Smith, M. S., Boston, M., & Hogel, M. (2008). Case stories: Supporting teachers reflection and collaboration on the implemention of cognitively challenging mathematical tasks. In *Inquiry into Mathematics Teacher Eduction* (pp. 71–83).

Huinker, D., & Bill, V. (2017). *Taking action: Implementing effective mathematics teaching practices K-grade 5*. (M. S. Smith, Ed.). Reston, VA: National Council of Teachers of Mathematics.

Huinker, D., & Schefelker, B. The Case of Ms . Bouchard and the Hungry Caterpillar Task (2016). Retrieved from https://www.nctm.org/Conferences-and-Professional-Development/Principles-to-Actions-Toolkit/The-Case-of-Ms_-Bouchard-and-the-Hungry-Caterpillar-Task/

Hull, T. H., Balka, D. S., & Miles, R. H. (2009). *A Guide to Mathematics Coaching: Processes for*

Increasing Student Achievement. Thousand Oaks, CA: Corwin Press, Inc.

Huntley, M. A., Rasmussen, C. L., Villarubi, R. S., Sangtong, J., & Fey, J. T. (2000). Effects of Standards-Based Mathematics Education: A Study of the Core-plus Mathematics Project Algebra and Functions Strand. *Journal for Research in Mathematics Education*, *31*(3), 328–361. https://doi.org/10.2307/749810

Illustrative Mathematics. (2016a). 1.NBT Ford and Logan Add 45+36. Retrieved from https://www.illustrativemathematics.org/content-standards/tasks/2068

Illustrative Mathematics. (2016b). 5 . NF Scaling Up and Down. Retrieved August 20, 2002, from https://www.illustrativemathematics.org/content-standards/5/NF/B/5/tasks/2101

Joyce, B., & Showers, B. (1980). Improving inservice training: The messages of research. *Educational Leadership*, (February), 379–385.

Joyce, B., & Showers, B. (1982). The coaching of teaching. *Educational Leadership*, (October), 4–10.

Joyce, B., & Showers, B. (2002). Designing training and peer coaching. In *Student Achievement Through Staff Development* (3rd ed., pp. 69–94). Alexandria, VA: Association for Supervision and Curriculum Development.

Kennedy, M. M. (1999). Form and substance in mathematics and science professional development. *NISE Brief*. Retrieved from http://www.eric.ed.gov/ERICWebPortal/recordDetail?accno=ED435552

Kennedy, M. M. (2016). How does professional development improve teaching? *Review of Educational Research*, 1–36. https://doi.org/10.3102/0034654315626800

Knapp, N. F., & Peterson, P. L. (1995). Teachers' Interpretations of "CGI" after Four Years: Meanings and Practices. *Journal for Research in Mathematics Education*, *26*(1), 40–65. https://doi.org/10.2307/749227

Kohler, F. W., Crilley, K. M., Shearer, D. D., & Good, G. (1997). Effects of peer coaching on teacher and student outcomes. *The Journal of Educational Research*, *90*(4), 240–250.

Kraft, M. A., Blazar, D., & Hogan, D. (2018). The Effect of Teacher Coaching on Instruction and Achievement: A Meta-Analysis of the Causal Evidence. *Review of Educational Research*, *88*(4), 547–588. https://doi.org/10.3102/0034654318759268

Krupa, E. E., & Confrey, J. (2012). Using instructional coaches: Customized professional development in an integrated high school mathematics program. In J. M. Bay-Williams & W. R. Speer (Eds.), *Professional Collaborations in Mathematics Teaching and Learning: Seeking Success for All, Seventy-fourth Yearbook* (pp. 161–174). Reston, VA: National Council of Teachers of Mathematics.

Lampert, M., Beasley, H., Ghousseini, H., Kazemi, E., & Franke, M. L. (2010). Using Designed

Instructional Activities to Enable Novices to Manage Ambitious Mathematics Teaching. In M. K. Stein & L. Kucan (Eds.), *Instructional Explanations in the Disciplines* (pp. 129–141). Springer US. https://doi.org/10.1007/978-1-4419-0594-9_9

Lampert, M., Boerst, T. A., & Graziani, F. (2011). Organizational resources in the service of school-wide ambitious teaching practice. *Teachers College Record*, *113*(7), 1361–1400. Retrieved from http://www.tcrecord.org/DefaultFiles/SendFileToPublic.asp?ft=pdf&FilePath=c:%5CWebSites%5Cwww_tcrecord_org_documents%5C38_16072.pdf&fid=38_16072&aid=2&RID=16072&pf=Content.asp?ContentID=16072

Lampert, M., Franke, M. L., Kazemi, E., Ghousseini, H., Turrou, A. C., Beasley, H., … Crowe, K. (2013). Keeping it complex: Using rehearsals to support novice teacher learning of ambitious teaching. *Journal of Teacher Education*, *64*(3), 226–243. https://doi.org/10.1177/0022487112473837

Learning Mathematics for Teaching Project. (2011). Measuring the mathematical quality of instruction. *Journal of Mathematics Teacher Education*, *14*(1), 25–47. https://doi.org/10.1007/s10857-010-9140-1

Lester Jr., F. K. (Ed.). (2007). *Second Handbook of Research on Mathematics Teaching and Learning*. Charlotte, NC: Information Age Publishing.

Loucks-Horsley, S., Stiles, K. E., Mundry, S., Love, N., & Hewson, P. W. (2012). *Designing Professional Development for Teachers of Science and Mathematics: Third Edition*. *Designing Professional Development for Teachers of Science and Mathematics* (3rd ed.). Thousand Oaks, CA: Corwin Press. https://doi.org/10.4135/9781452219103

Math Design Collaborative. (2015). PA – MDC GRADE 3: Perimeter and Area, Problem solving formative assessment lesson. Math Design Collaborative.

Matsumura, L. C., Garnier, H. E., Correnti, R., Junker, B., & Bickel, D. D. (2010). Investigating the effectiveness of a comprehensive literacy coaching program in schools with high teacher mobility. *The Elementary School Journal*, *111*(1), 35–62. https://doi.org/10.1086/653469

Matsumura, L. C., Garnier, H. E., Slater, S. C., & Boston, M. (2008). Toward measuring instructional interactions at-scale. *Educational Assessment*, *13*(4), 267–300. https://doi.org/10.1080/10627190802602541

Matsumura, L. C., Garnier, H. E., & Spybrook, J. (2013). Literacy coaching to improve student reading achievement: A multi-level mediation model. *Learning and Instruction*, *25*, 35–48. https://doi.org/10.1016/j.learninstruc.2012.11.001

Michaels, S., O'Connor, C., & Resnick, L. B. (2008). Deliberative discourse idealized and realized: Accountable talk in the classroom and in civic life. *Studies in Philosophy and Education*, *27*(4), 283–297. https://doi.org/10.1007/s11217-007-9071-1

Michigan State University College of Natural Sciences. (n.d.). CMP Lesson Structure - Launch, Explore, Summarize.

Mindset Works Inc. (2016). You Can Grow Your Intelligence. *Brainology*. Retrieved from http://www.nais.org/publications/ismagazinearticle.cfm?itemNumber=150439

Morse, A. (2009). *Cultivating a Math Coaching Practice: A Guide for K-8 Math Educators*. Thousand Oaks, CA: Corwin Press, Inc.

Murray, S., Ma, X., & Mazur, J. (2009). Effects of Peer Coaching on Teachers' Collaborative Interactions and Students' Mathematics Achievement. *The Journal of Educational Research*, *102*(3), 203–212. Retrieved from http://pitt.idm.oclc.org/login?url=http://search.proquest.com/docview/204194955?accountid=14709

National Center for Education Statistics [NCES]. (2003). Teaching mathematics in seven countries: Results from the TIMSS 1999 video study. *Education Statistics Quarterly*, *5*(1), 7–15. https://doi.org/Washington, DC

National Council of Teachers of Mathematics. (1989). *Curriculum and Evaluation Standards for School Mathematics*. Reston, VA: National Council of Teachers of Mathematics.

National Council of Teachers of Mathematics. (1991). *Professional Standards for Teaching Mathematics*. Reston, VA: National Council of Teachers of Mathematics.

National Council of Teachers of Mathematics. (1995). *Assessment Standards for School Mathematics*. Reston, VA: National Council of Teachers of Mathematics.

National Council of Teachers of Mathematics. (2000a). *Principles and Standard for School Mathematics*. Reston, VA: National Council of Teachers of Mathematics.

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

National Council of Teachers of Mathematics. (2014). *Principles to action: Ensuring mathematical success for all*. Reston, VA: National Council of Teachers of Mathematics.

National Governors Association Center for Best Practices and the Council of Chief State School Officers. (2010). Common Core State Standards. Washington, D.C.: National Governors Association Center for Best Practices, Council of Chief State School Officers.

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

National Research Council. (2005). *How Students Learn: History, Mathematics, and Science in the Classroom*. (M. S. Donovan & J. D. Bransford, Eds.). Washington, DC: National Academies Press.

Neufeld, B., & Roper, D. (2003). *Coaching: A Strategy for Developing Instructional Capacity*.

Cambridge, Mass.

NGSS Lead States. (2013). *Next Generation Science Standards: For states, by states*. Washington, DC: National Academies Press. https://doi.org/10.17226/18290

Parrish, S. D. (2016). Number talks build numerical reasoning. *Teaching Children Mathematics*, *18*(3), 198. https://doi.org/10.5951/teacchilmath.18.3.0198

Polly, D. (2012). Supporting mathematics instruction with an expert coaching model. *Mathematics Teacher Education and Development*, *14*, 78–93.

Putnam, R. T., & Borko, H. (2000). What Do New Views of Knowledge and Thinking Have to Say About Research on Teacher Learning? *Educational Researcher*, *29*(1), 4–15. https://doi.org/10.3102/0013189X029001004

Removcik, J. A. (2014). *Long-term use of standards-based instructional practices in mathematics classrooms in Southwestern Pennsylvania*. Retrieved from http://search.proquest.com.ezp.waldenulibrary.org/docview/1609203637/abstract/4D9A48AB76714C27PQ/1?accountid=14872%5Cnhttp://media.proquest.com.ezp.waldenulibrary.org/media/pq/classic/doc/3455073681/fmt/ai/rep/NPDF?hl=high,instructional,practises,tsarevit

Resnick, L. B., & Hall, M. W. (1998). Learning organizations for sustainable education reform. *Daedalus*, *127*(4), 89–116.

Resnick, L. B., Matsumura, L. C., & Junker, B. (2006). *Measuring Reading Comprehension and Mathematics Instruction in Urban Middle Schools : A Pilot Study of the Instructional Quality Assessment* (Vol. 1522). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.

Resnick, L. B., Stein, M. K., & Coon, S. (2008). Standards-Based Reform: A Powerful Idea Unmoored. In R. Kahlenbert (Ed.), *Improving on No Child Left Behind* (pp. 103–138). New York, NY: The Century Foundation.

Romero, S. M., & Winters, D. M. (2013). *Report of PSSA and Keystone Exam Data Analysis by*. Pittsburgh, PA.

Russell, J. L., Correnti, R., Stein, M. K., Bill, V., Hannan, M., Schwartz, N., … Matthis, C. (2019). Learning From Adaptation to Support Instructional Improvement at Scale: Understanding Coach Adaptation in the TN Mathematics Coaching Project. *American Educational Research Journal*, *XX*(X), 000283121985405. https://doi.org/10.3102/0002831219854050

Russell, J. L., Correnti, R., Stein, M. K., Hannan, M., & Bill, V. (2017). Adaptive Integration-March2017-v2 (1).

Sawada, D., Piburn, M. D., Judson, E., Turley, J., Falconer, K., Benford, R., & Bloom, I. (2002). Measuring Reform Practices in Science and Mathematics Classrooms: The Reformed Teaching Observation Protocol. *School Science and Mathematics*, *102*(6), 245–253. https://doi.org/10.1111/j.1949-8594.2002.tb17883.x

Schifter, D. E., & O'Brien, D. C. (1997). Interpreting the standards: translating principles into practice. *Teaching Children Mathematics*, *4*(4), 202–206. Retrieved from http://go.galegroup.com.pitt.idm.oclc.org/ps/i.do?p=AONE&u=upitt_main&id=GALE%7CA20435260&v=2.1&it=r&sid=summon

Showers, B. (1984). *Peer coaching: A strategy for facilitating transfer of training. A CEPM R&D Report.* Eugene, OR.

Showers, B. (1985). Teachers coaching teachers. *Educational Leadership*, *43*(April), 19–27.

Showers, B., & Joyce, B. (1996). The evolution of peer coaching. *Educational Leadership*, *12*(March), 1–5.

Shulman, L. S. (1986). Those who Understand : Knowledge Growth in Teaching. *Educational Researcher*, *15*(2), 4–14. https://doi.org/http://www.jstor.org/stable/1175860

Shulman, L. S. (1987). Knowledge and Teaching: Foundations of the New Reform. *Harvard Educational Review*, *57*(1), 1–21. https://doi.org/10.1007/SpringerReference_17273

Shulman, L. S. (1998). Theory, practice, and the education of professionals. *The Elementary School Journal*, *98*(5), 511–526. Retrieved from https://search-proquest-com.pitt.idm.oclc.org/docview/224520352?pq-origsite=summon

Silver, E. A. (1994). On Mathematical Problem Posing. *Educational Studies in Mathematics*, *1*(14), 19–28.

Silver, E. A., Smith, M. S., & Nelson, B. S. (1995). The QUASAR Project: Equity concerns meet mathematics reform in the middle school. In W. G. Secada, E. Fennema, & L. Byrd (Eds.), *New Directions for Equity in Mathematics Education* (pp. 9–39). New York, NY: Cambridge University Press.

Silver, E. A., & Stein, M. K. (1996). The Quasar Project. *Urban Education*, *30*(4), 476–521. https://doi.org/10.1177/0042085996030004006

Smith, M. S. (2001). *Practice-Based Professional Development for Teachers of Mathematics*. Reston, VA: National Council of Teachers of Mathematics.

Smith, M. S. (Ed.). (2017). *Taking action: Implementing effective mathematics teaching practices*. Reston, VA: National Council of Teachers of Mathematics.

Smith, M. S., Boston, M., & Huinker, D. (2017). Preface. In M. S. Smith (Ed.), *Taking Action: Implementing Effective Mathematics Teaching Practices* (pp. v–vi). Reston, VA: National Council of Teachers of Mathematics, Inc.

Smith, M. S., Hughes, E. K., Engle, R. A., & Stein, M. K. (2009). Orchestrating Discussions. *Mathematics Teaching in the Middle School*, *14*(9), 548–556. Retrieved from http://www.jstor.org.pitt.idm.oclc.org/stable/41182933

Smith, M. S., Steele, M., & Raith, M. L. (2017). *Taking action: Implementing effective mathematics teaching practices in grades 6-8*. Reston, VA: National Council of Teachers of Mathematics.

Smith, M. S., & Stein, M. K. (2011). *5 Practices for Orchestrating Productive Mathematics Discussions*. Reston, VA: National Council of Teachers of Mathematics.

Smith, M. S., & Stein, M. K. (2018). *5 practices for orchestrating productive mathematics discussions; Second Edition* (Second). Reston, VA: National Council of Teachers of Mathematics.

Stanford Graduate School of Education. (n.d.). youcubed Brain Science.

Staples, M. (2007). Supporting Whole-class Collaborative Inquiry in a Secondary Mathematics Classroom. *Cognition and Instruction*, *25*(2–3), 161–217. https://doi.org/10.1080/07370000701301125

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. https://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*, 50–80.

Stein, M. K., Remillard, J. T., & Smith, M. S. (2007). How curriculum influences student learning. In F. K. Lester, Jr. (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 319–369). Charlotte, NC: Information Age Publishing.

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*(4), 268–275.

Stein, M. K., Smith, M. S., Henningsen, M., & Silver, E. A. (2000). *Implementing Standards-based Mathematics Instruction : A Casebook for Professional Development*. New York: Teachers College Press. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,uid&db=nlebk&AN=34555&scope=site

Stein, M. K., Smith, M. S., Henningsen, M., & Silver, E. A. (2009). *Implementing standards-based mathematics instruction: A casebook for professional development* (Second Edi). New York, NY and Reston, VA: Teachers College Press and National Council of Teachers of Mathematics.

Stein, M. K., & Wang, M. C. (1988). Teacher development and school improvement: The process of teacher change. *Teaching and Teacher Education*, *4*(2), 171–187. https://doi.org/10.1016/0742-051X(88)90016-9

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

Stigler, J. W., & Hiebert, J. (2004). Improving mathematics teaching. *Educational Leadership*, *61*, 12–16.

Tarr, J. E., Reys, R. E., Reys, B. J., Chavez, O., Shih, J., & Osterlind, S. (2008). The impact of middle-grades mathematics curricula and the classroom learning environment on student achievement. *Journal for Research in Mathematics Education*, *39*(3), 247–280.

The Charles A. Dana Center at the University of Texas at Austin. (n.d.). Maria's Money: Student Work. Retrieved from http://www.insidemathematics.org/assets/classroom-videos/public-lessons/4th-grade-number-operations-multiplication-division/student-sample-set-1.pdf

The National Commission on Excellence in Education. (1983). *A Nation at Risk: The Imperative for Educational Reform*. Washington, D.C. Retrieved from https://www.edreform.com/wp-content/uploads/2013/02/A_Nation_At_Risk_1983.pdf

The work of teaching. (n.d.). Retrieved from http://www.teachingworks.org/

U.S. Department of Education. (2008). *The Final Report of the National Mathematics Advisory Panel*. Washington, D.C.

University of Pittsburgh (Collaborative for Evaluation and Assessment Capacity). (2016). MSP-MSC 2016-2017 Pre-Survey.

Wang, Y. (2013). *Southwestern Pennsylvania Assessment of Teacher Content Knowledge Year- - - End Report*. Pittsburgh.

Wang, Y., & Romero, S. M. (2013). *Report of Math and Science Educator Survey Analysis*. Pittsburgh, PA.

Warfield, J., Wood, T., & Lehman, J. D. (2005). Autonomy, beliefs and the learning of elementary mathematics teachers. *Teaching and Teacher Education*, *21*(4), 439–456. https://doi.org/10.1016/j.tate.2005.01.011

Weiss, I. R., Pasley, J. D., Smith, P. S., Banilower, E. R., & Heck, D. J. (2003). *Looking Inside the Classroom: K12 Mathematics and Science Education in the United States*. *SSRN Electronic Journal*. Chapel Hill, NC. https://doi.org/10.2139/ssrn.156634

West, L., & Staub, F. C. (2003). *Content-Focused Coaching: Transforming Mathematics Lessons*. Portsmouth, NH: Heinemann.

What is ambitious teaching. (n.d.).

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. https://doi.org/10.5951/jresematheduc.46.3.0270

Wiliam, D. (2007). Research brief. Five “key strategies” for effective formative assessment. https://doi.org/10.1016/S2213-8587(14)70016-6

Wiliam, D. (2013). Assessment for Learning. Retrieved March 21, 2019, from https://www.youtube.com/watch?v=jOHhJorBjwU

Windschitl, M., Thompson, J., Braaten, M., & Stroupe, D. (2012). Proposing a core set of instructional practices and tools for teachers of science. *Science Education*, *96*(5), 878–903. https://doi.org/10.1002/sce.21027

Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education*, *11*(1), 57–67. Retrieved from https://search-proquest-com.pitt.idm.oclc.org/docview/763045390?pq-origsite=summon