DEVELOPING SECONDARY MATHEMATICS TEACHERS' KNOWLEDGE OF AND CAPACITY TO IMPLEMENT INSTRUCTIONAL TASKS WITH HIGH LEVEL COGNITIVE DEMANDS

by

Melissa D. Boston

BS, Mathematics Education, Grove City College, 1992

MA, Mathematics, University of Pittsburgh, 1997

Submitted to the Graduate Faculty of the

School of Education in partial fulfillment

of the requirements for the degree of

Educational Doctorate in Mathematics Education

University of Pittsburgh

2006

UNIVERSITY OF PITTSBURGH

School of Education
Department of Instruction and Learning

This dissertation was presented

by

Melissa D. Boston

It was defended on

April 4, 2006

and approved by

Dr. Ellen Ansell

Dr. Gaea Leinhardt

Dr. Mary Kay Stein

Dissertation Director: Dr. Margaret S. Smith

DEVELOPING SECONDARY MATHEMATICS TEACHERS' KNOWLEDGE OF AND CAPACITY TO IMPLEMENT INSTRUCTIONAL TASKS WITH HIGH LEVEL COGNITIVE DEMANDS

Melissa D. Boston, EdD

University of Pittsburgh, 2006

This study analyzed mathematics teachers' selection and implementation of instructional tasks in their own classrooms before, during, and after their participation in a professional development workshop focused on the cognitive demands of mathematical tasks. Eighteen secondary mathematics teachers participated in a six-session professional development workshop under the auspices of the Enhancing Secondary Mathematics Teacher Preparation (ESP) Project throughout the 2004-2005 school year. Data collected from the ESP workshop included written artifacts created during the professional development sessions and videotapes of each session. Data collected from teachers included a pre/post measure of teachers' knowledge of the cognitive demands of mathematical tasks, collections of tasks and student work from teachers' classrooms, lesson observations, and interviews. Ten secondary mathematics teachers who did not participate in the ESP workshop served as the contrast group, completed the pre/post measure, and participated in one lesson observation.

Analysis of the data indicated that the ESP workshop provided learning experiences for teachers that transformed their previous knowledge and instructional practices. ESP teachers enhanced their knowledge of the cognitive demands of mathematical tasks; specifically, they improved their ability to identify and describe the characteristics of tasks that influence students' opportunities for learning. Following their participation in ESP, teachers were more frequently

selecting high-level tasks as the main instructional tasks in their own classrooms. ESP teachers also improved their ability to maintain high-level cognitive demands during implementation. Student work implementation significantly improved from Fall to Spring, and comparisons of the implementation of high-level student work tasks indicated that high-level demands were less likely to decline in Spring than in Fall. Lesson observations did not yield statistically significant results from Fall to Spring; however, significant differences existed between ESP teachers and the contrast group in task selection and implementation during lesson observations. ESP teachers also outperformed the contrast group on the post-measure of the knowledge of cognitive demands of mathematical tasks. None of the significant differences were influenced by the use of a reform vs. traditional curricula in teachers' classrooms. Teachers who exhibited greater improvements more frequent contributions and more comments on issues of implementation than teachers who exhibited less improvement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Acknowledgements

Many people have provided assistance throughout the long and fruitful journey that has culminated in the completion of my dissertation. I extend my deepest gratitude to my research advisor and mentor, Dr. Margaret (Peg) Smith. Every meeting with you enhanced my own learning and professional growth. Thank-you for providing countless opportunities for me to grow as a teacher-educator, and thank-you for seeing something in me that was worth your time, effort, patience, and perseverance over this process.

I would like to thank my dissertation committee for their advice, direction, guidance, feedback, and most importantly, for the accomplishments in their prior research that contributed to this investigation. To Dr. Gaea Leinhardt, thank you for providing opportunities early in the dissertation process that enabled me to move forward. To Dr. Mary Kay Stein, I am indebted to your prior research on QUASAR that enhanced this investigation in numerous ways. To Drs. Ellen Ansell and Peg Smith, thank you for creating the ESP project and for providing many learning opportunities throughout my (extended) time at Pitt as we planned and discussed the mathematics methods courses and the ESP workshop.

Next, I would like to thank the teachers who participated in ESP. Your active engagement throughout the ESP workshop and your willingness to provide data from your classrooms made this study possible. Your commitment to teaching mathematics, measured in part by your willingness to devote Saturdays to your own learning and professional growth, is admirable. I am also indebted to the ESP professional development team: Ellen Ansell, Peg Smith, Michael Hogel, Elizabeth Hughes, Mary Lou Metz, and Michael Steele. The opportunity to collaborate with you all in planning and presenting the ESP sessions was a wonderful experience. The success of the ESP workshop can be attributed to your insights into teacher-learning and your

commitment to designing quality professional development experiences. The success of the ESP workshop can also be attributed to the efforts of Kathy Day, who supported the ESP project and its team members in countless ways. To Kathy, the thoughtfulness and thoroughness you apply to everything you do, and the kindness and encouragement you have consistently provided, will never be forgotten.

I also extend my appreciation to fellow graduate students Amy Hillen, Elizabeth Hughes, Mary Lou Metz, Jennifer Mossgrove, and Michael Steele for providing assistance, encouragement, and much needed lunch breaks (and e-mail) throughout the dissertation process. I thank Michael Steele and Jennifer Mossgrove for assisting with data collection and coding and Elaine Rubenstein and Richard Lowry (creator of the Vassarstats website) for providing statistical support.

I also thank my parents, Keith and Linda Patterson, and my mother-in-law, Georgette Dorsey. Their support and assistance in providing child-care (and child-entertainment) enabled me to pursue my doctoral studies and to devote time toward completing my "paper" without feeling (too) guilty.

My sincerest appreciation and deepest love are extended to my husband, David Boston, whose encouragement, love, and support helped me to make progress (even when I didn't want to); who didn't object to being the only father at Brownies or dance lessons; and who managed the responsibilities of parenthood and running a home in the final stretch so that I could completely and totally focus on completing my "paper."

Finally, I thank my daughters, Lauren and Lindsey Boston, for their over-abundant love, artwork, hugs and kisses. Someday you will understand why this was worth doing.

My "paper" is finally finished!!

1. CHAPTER 1: THE RESEARCH PROBLEM

1.1. Introduction

Over a decade ago, the National Council of Teacher of Mathematics (NCTM) unveiled standards for the teaching and learning of mathematics asserting the importance of mathematical thinking, reasoning, and understanding in the lives and futures of American students (NCTM, 1989) and portraying a vision of the type of mathematics teaching necessary to attain this goal (NCTM, 1991). This vision differed in significant ways from the pervasive form of mathematics instruction at the time – one in which the teacher's role consisted of presenting isolated facts and procedures for students to master and the students' role consisted of memorizing and accurately reproducing these facts and procedures (NCTM, 1991; Romberg & Carpenter, 1986; Putnam, Lampert & Peterson, 1990). NCTM called for dramatic shifts in the roles of both teachers and students in order for students to develop the ability to think, reason, conjecture, and communicate mathematically. Consistent with constructivist perspectives on learning mathematics with understanding (cf. Hiebert & Carpenter, 1992; Putnam, et al., 1990; Voigt, 1994), the role of students was changed from passive recipients of facts and procedures to active constructers of mathematical knowledge. Teachers were now to serve as facilitators of students' learning by providing classroom experiences in which students can engage with rich mathematical tasks, develop connections between mathematical ideas and between different

representations of mathematical ideas, and collaboratively construct and communicate their mathematical thinking.

Unfortunately, these new roles for teachers and students have yet to be realized in the majority of mathematics classrooms in the United States. As recently as 1999, an analysis of a random sample of 100 U.S. mathematics classrooms participating in the TIMSS Video Study (USDE-NCES, 2003) indicated that U.S. students spend virtually all of their time in mathematics class performing procedures, stating concepts, or simply providing answers. Less than 1% of students' mathematical experiences involved making connections, "constructing relationships between facts, ideas, or procedures, …or engaging in mathematical reasoning such as conjecturing, generalizing, and verifying" (p. 98), and over half of mathematics instructional time was spent reviewing previously learned concepts or procedures in ways that did not advance the mathematical ideas (USDE-NCES, 2003). Hence, efforts to reform mathematics teaching and learning have not resulted in the wide-scale implementation of the type of mathematics instruction that fosters learning mathematics with understanding.

The absence of reform-oriented features from mathematics classrooms would not be a concern if American students exhibited strong performance on international, national or state assessments of mathematical ability. However, U.S. students continue to post substandard performance on mathematical assessments at all levels. The mean performance of U.S. students on the mathematics portion of the 1995 TIMSS was significantly lower than the international average in both 8th and 11th-grades (USDE-NCES, 2003). Results for the 2003 National Assessment of Educational Progress (NAEP) indicate that only 29% of the national sample of eighth-grade students demonstrated mathematical proficiency (NCES, 2004). Over half of students in grades 8 and 11 fell below the mathematically proficient level on the 2003

Pennsylvania System of School Assessment (PSSA) (PDE, 2004). Consistent indications that our students are not able to engage with or understand mathematics in ways that are deemed valuable to their future success are cause for great concern.

In *Before Its Too Late,* the National Commission on Mathematics and Science Teaching (NCMST) for the 21st Century (USDE, 2000) acknowledged the current, persistent need for improved achievement of American students in mathematics, and asserted, "The most direct route to improving mathematics and science achievement for all students is better mathematics and science teaching. In other words, better teaching is the lever for change (p.18)." The type of improvements in current teaching practices will require *transformative* learning on the part of teachers – learning that will catalyze changes in teachers' long-held, underlying beliefs about effective teaching and learning of mathematics (Thompson & Zeuli, 1999). Hence, professional development experiences that provide opportunities for mathematics teachers to reconsider their current instructional practices in light of new ideas and experiences can serve as a vehicle for moving teachers toward new practices that improve students' learning of mathematics.

The purpose of this study is to determine the extent to which professional development that focuses on the cognitive demands of mathematical tasks and the maintenance of high-level cognitive demands throughout an instructional episode will influence the ways in which mathematics teachers select and implement the tasks they use in their own classrooms. The guiding hypothesis is that providing teachers with experiences in analyzing the cognitive demands of mathematical tasks and the implementation of cognitively challenging tasks will enable teachers to select and implement mathematical tasks in their own classrooms in ways that provide increased opportunities for students' learning. The remainder of this chapter will justify using task selection and implementation as a focal point for improving mathematics teaching.

1.2. Background

The assertion that improvements in mathematics teaching will generate improvements in students' learning rests on the assumption that teaching influences learning. According to Doyle (1988), "the work students do, defined in large measure by the tasks teachers assign, determines how they think about a curricular domain and come to understand its meaning (p.167)." He proposes that teachers influence students' learning of mathematics by structuring the academic work that students do during mathematics class. Teachers select the tasks that constitute students' mathematical work, and teachers specify what processes can be used, what resources are available, what products are expected, and how those products will be evaluated. Teachers set the parameters for the selection and implementation of instructional tasks, and instructional tasks set the parameters for what students have an opportunity to learn. Hence, improving students' learning can be accomplished by improving the tasks teachers select and implement in their classrooms (Doyle, 1983, 1988; Stein & Lane, 1996; Hiebert & Wearne, 1993). In this study, I will argue that one route to improving mathematics teaching in ways that will improve students' learning is to engage mathematics teachers in professional development experiences that convey the significance of mathematical tasks in influencing students' opportunities to learn mathematics. If mathematical tasks can serve as a catalyst for learning mathematics with understanding, then professional development for mathematics teachers should explicate the importance of the mathematical tasks they select for instruction and the ways in which these tasks are implemented in their classrooms. The following sections will discuss the influence of tasks on students' opportunities to learn mathematics with understanding and thus establish the importance of making task research a salient aspect of professional development for mathematics teachers.

1.2.1. The Influence of Tasks on Students' Learning

A ***mathematical task*** is defined as a set of problems or single complex problem that focuses students' attention on a particular mathematical idea (Stein, Grover, & Henningsen, 1996). The construct of task also encompasses the intellectual and physical products that are expected of students, the operations that students are to use to obtain the desired products, and the resources that are available to assist students in completing these products (Doyle, 1983). On a practical level, tasks influence student learning because working on mathematical tasks constitutes what students *do* during the majority of their time in mathematics class. Students in all seven countries analyzed in the TIMSS Video Study (USDE-NCES, 2003) (i.e., Australia, Czech Republic, Hong Kong, Japan, Netherlands, Switzerland, and the United States) spent over 80% of their time in mathematics class working on mathematical tasks. On a theoretical level, Doyle offers two premises for why "tasks form the basic treatment unit in classrooms" (1983, p. 162). First, a mathematical task draws students' attention toward a particular mathematical concept and provides certain information surrounding that concept (Doyle, 1983). Students are exposed to (and thus have an opportunity to learn) the concepts embedded in the tasks they complete. Students are not exposed to (and thus have much less of an opportunity to learn) content that is not represented in the tasks they complete. Tasks thus influence learning by defining the mathematical content that students have an opportunity to learn.

Second, tasks influence student learning by setting parameters for the ways in which information about the mathematical concept can be operated on or processed (Doyle, 1983). Students will become skilled at what they have an opportunity to actually do in mathematics class. If students' academic work consists of practicing procedural computations, they are likely

to become facile with computational skills; however, if students spend their time reflecting on why things work the way they do, how ideas are connected to their prior knowledge, or how ideas and procedures compare and contrast, then they are likely to be constructing new relationships and new understandings of mathematics (Hiebert, Carpenter, Fennema, Fuson, Wearne, Murray, Olivier, & Human, 1997). Given that students are not likely to spontaneously do more than what a task requires, Doyle (1983) notes that students will identify the information and operations that are necessary to accomplish the task and will adjust their strategies for working with that information depending on what they perceive is required by the task. If the tasks students typically encounter in mathematics class lead them to perceive mathematical work as imitating a prescribed procedure, executing the steps in an algorithm, or stating memorized facts, students will adjust their strategies for working on mathematical tasks to correspond with these perceptions. However, if the tasks in which students engage set norms that require extended thinking, reasoning, problem-solving and communication, then these strategies will become part of students' repertoire for performing mathematical work. The collection of tasks students perform throughout their mathematics instruction form students' notions of how work in mathematics is performed and of what mathematics *is* in general (Doyle, 1988).

Hence, different types of tasks provide different opportunities for students' learning and place different expectations on students' thinking. A task that entails only memorization will result in much different learning than a task that requires problem-solving, conjecturing, and reasoning (Smith & Stein, 1998). The thinking processes that a task has the potential to elicit have been referred to as the ***cognitive demands*** of an instructional task (Stein, Grover, & Henningsen, 1996). Tasks with high-level cognitive demands have the potential to engage students with high-level thinking processes, such as problem-solving, conjecturing, justifying,

generalizing, or proving (Van de Walle, 2004). Mathematical tasks with high-level cognitive demands (hereafter referred to as ***high-level tasks***) provide opportunities for reflection on and communication of important mathematics, where the mathematics in the task is intellectually challenging for students, connects with students' prior knowledge, and leaves behind valuable mathematical "residue" (Hiebert, et al., 1997, p. 18). High-level tasks often involve mathematizing, or describing a situation in terms of its quantitative relationships (Putnam, et al., 1990). Tasks described as "worthwhile tasks" by NCTM (2000) are examples of tasks with high-level cognitive demands.

Stein and colleagues (Stein, et al., 1996) identify two categories of high-level mathematical tasks: (1) "doing mathematics" -- tasks that require complex and non-algorithmic thinking, require students to draw on prior knowledge to solve the task, and provide opportunities for students to explore and to understand the nature of mathematical concepts, processes, or relationships; and (2) "procedures with connections" – tasks that engage students in using broad general procedures for the purpose of developing deep understandings of the underlying mathematical concepts or processes, and often require students to form connections between multiple mathematical representations. A distinguishing feature between Stein et al.'s two categories of high-level tasks is that "doing mathematics" tasks do not provide or imply a pathway for solving the task (and thus provide opportunities for students to develop their own problem solving strategies), where a "procedures with connections" task often states or implies a pathway for students to follow (though this pathway requires students to make sense of mathematical ideas at each step along the way).

In contrast, tasks with low-level cognitive demands have the potential to engage student with low-level thinking processes, such as memorization or the application of procedures with no

connection to meaning or understanding (Stein, et al., 1996; Doyle, 1983). Tasks with low-level cognitive demands (hereafter referred to as ***low-level tasks***) are not inappropriate or "bad" instructional tasks. If the goal of an instructional episode is for students to memorize formulae, reproduce a demonstrated example, or practice a given procedure, then tasks that require low levels of cognitive demand are appropriate. However, if the goal of an instructional episode is for students to think, reason, and engage in problem-solving, then instruction must be based on high-level mathematical tasks (Stein & Lane, 1996; NCTM, 2000; Hiebert, et al., 1997). This assertion is supported by social-constructivist theories of students' learning, which contend that mathematical understanding develops through the process of 1) building on students' current knowledge; 2) promoting active engagement with and reflection on new mathematical ideas or experiences in ways that generate a re-organization of students' current knowledge (Romberg & Carpenter, 1986); and 3) providing opportunities for collaborative communication, negotiation, and construction of the new mathematical ideas (Voigt, 1994; Cobb, Yackel, & Wood, 1991). A characteristic of mathematical understanding is the ability to recognize a mathematical idea within a variety of representations, to work with the idea within a specific representation, and to translate the idea between different representations (Lesh, Post, & Behr, 1987). Hence, mathematical tasks should encourage students to represent and structure mathematical ideas, both physically and mentally, in ways that facilitate connections between concepts, facts, and procedures (Hiebert & Carpenter, 1992; Greeno, 1991).

In summary, theories of students' learning of mathematics support the contention that tasks with high-level cognitive demands are likely to foster mathematical understanding (Hiebert & Carpenter, 1992; Putnam, et al., 1990). Empirical research to support this claim will be provided in the following section.

1.2.2. The Relationship between High-Level Tasks and Student Learning

Results from a growing body of research support the positive influence of high-level tasks on students' learning of mathematics. This research indicates that curricular materials specifically developed to contain high-level tasks (USDE, 1999) are successful at improving students' performance on state and national tests of mathematical achievement (e.g., Fuson, Carroll, & Druek, 2000; Riordan & Noyce, 2001; Schoen, Fey, Hirsch, & Coxford, 1999), at improving students' understanding of important mathematical concepts (e.g., Ben-Chaim, Fey, Fitzgerald, Benedetto, & Miller, 1998; Huntley, Rasmussen, Villarubi, Sangtong, & Fey, 2000; Thompson & Senk, 2001; Reys, Reys, Lappan, & Holliday, 2003), and at improving students abilities to reason, communicate, problem-solve and make mathematical connections (e.g., Ridgeway, Zawojewski, Hoover, & Lambdin, 2003; Schoenfeld, 2002). Engaging students with tasks that elicit high-level cognitive demands appears to have a positive effect on students' development of mathematical understanding.

In their analysis of mathematics lessons from the QUASAR Project[1], Stein and Lane (1996) further explicate the relationship between high-level tasks and student learning. Teachers in the QUASAR Project were attempting to transform their instructional practices in ways that incorporated high-level tasks and rich mathematical discussions. The success of this endeavor varied greatly across the different schools in the project, piquing Stein and Lane to assess the degree to which variations in the implementation of reform-oriented features of mathematics instruction could be linked to variations in students' learning. Stein and Lane found that

[1] The QUASAR (Quantitative Understanding: Amplifying Student Achievement and Reasoning) Project was a national reform project aimed at assisting schools in economically disadvantaged communities to develop middle school mathematics programs that emphasized thinking, reasoning, and problem-solving (Silver & Stein, 1996).

instruction characterized by the use of high-level tasks generated greater student learning gains than instruction characterized by the use of low-level mathematical tasks. With student learning gains defined as "increases in the quality of students' performance from one time period to another…elicited by an assessment instrument that was designed to measure outcomes in mathematics aligned with NCTM (1989)" (Stein & Lane, 1996, p. 60), Stein and Lane's results provide an instantiation of how mathematical tasks determine what students have an opportunity to learn (Doyle, 1988, 1983). Basing instruction on high-level mathematical tasks increased students' ability to engage in high-level mathematical processes, such as thinking, reasoning, and problem-solving. Stein and Lane concluded that providing students with opportunities to explore high-level mathematical tasks during instruction "confers greater benefit to students than does exposure to tasks that emphasize lower levels of cognitive thinking from the start" (1996, p. 74).

Unfortunately, a predominance of curricular materials currently in use in the U.S. do not expose students to high-level tasks. In the recent TIMSS Video Study (USDE-NCES, 2003), 83% of the tasks observed by TIMSS researchers involved low-level cognitive demands (as described by Stein, et al., 1996, or Doyle, 1988) such as stating concepts or applying procedures. Students' opportunities to form meaningful mathematical connections are further limited by mathematics curricular materials that lack coherence. The mathematics curricula described by both Doyle (1988) and TIMSS (USDE-NCES, 2003) consisted of a set of discrete, unrelated tasks that tended to emphasize executing procedures rather than understanding concepts or engaging in problem-solving. This type of curricular presentation often results in the teaching of mathematics as isolated facts, concepts, and procedures, and severely hinders students' opportunities to develop mathematical connections. Only 16% of the tasks in U.S. mathematics classrooms observed by TIMSS researchers were related mathematically to the previous

instructional task, while 68% were identified as purely repetitious of previous tasks (USDE-NCES, 2003). The predominance of repetitious, procedural tasks in U.S. mathematics classrooms leaves little room for tasks that require thinking reasoning, and problem-solving.

Even when high-level tasks are present in U.S. mathematics classrooms, students are not guaranteed opportunities to engage with high-level cognitive demands. Maintaining the complexity of high-level tasks is not a trivial endeavor and is often shaped by teachers' and students' beliefs about how mathematics is best taught and learned (Borko & Putnam, 1995; Remillard, 1999). Teachers and students accustomed to traditional, directive styles of teaching and routinized, procedural tasks experience conflict and discomfort with the ambiguity and struggle that often accompany high-level tasks (Smith, 1995; Clarke, 1997). In response to ambiguity or uncertainty on how to proceed, some students disengage with the task or press the teacher for step-by-step instructions (Romanagno, 1994; Henningsen & Stein, 1997), thereby reducing the cognitive demands of the task. Teachers in the QUASAR Project varied in their ability to maintain the cognitive demands of high-level tasks, and QUASAR researchers determined that the greatest student learning gains occurred in classrooms where high-level cognitive demands were consistently maintained throughout the instructional episode (Stein & Lane, 1996). This finding is consistent with results from the TIMSS 1999 Video Study (USDE-NCES, 2003): teachers in countries whose students outperformed U.S students implemented high-level tasks in ways that maintained high-level cognitive demands. While teachers in the U.S. used percentages of high-level tasks consistent with the percentages used in many higher-performing countries, the most striking and significant difference between the U.S. and higher performing countries in the study was the inability of U.S. teachers to maintain high-level cognitive demands during instruction. Opportunities to engage in high-level thinking and

reasoning throughout an instructional episode appear to improve students' learning of mathematics. Such opportunities currently do not exist in the majority of U.S. classrooms. The following section describes how professional development for teachers of mathematics might improve American students' opportunities to learn mathematics with understanding.

1.2.3. Professional Development Focused on the Selection and Implementation of High-Level Mathematical Tasks

Improving students' opportunities to learn mathematics with understanding requires improving the cognitive demands of the tasks used during mathematics instruction and increasing students' opportunities to engage with high-level cognitive demands throughout an instructional episode. This, in turn, requires that mathematics teachers 1) select high-level tasks for instruction and 2) implement high-level tasks in ways that support students' understanding. Empirical evidence on the low percentage of instructional tasks with high-level cognitive demands and on the fate of tasks that began as high-level (USDE-NCES, 2003; Stein & Lane, 1996) indicates that these requirements are not typically present in mathematics classrooms. Hence, increasing students' exposure to and sustained engagement with high-level cognitive demands will require changes in the knowledge and instructional practices of mathematics teachers. In this investigation, changes in teachers' knowledge and instructional practices are anticipated to result from teachers' participation in professional development experiences specifically focused on the selection and implementation of high-level tasks.

Several professional development studies have linked teachers' participation in professional development to changes in teachers practice. For example, teachers participating in the Cognitively Guided Instruction (CGI) project began to make changes in their classrooms

consistent with the principles of CGI, such as implementing tasks from the workshops, listening to students' thinking, and building on students' mathematical thinking to make their instructional decisions (Fennema, Carpenter, Franke, Levi, Jacobs, Empson, 1996; Franke, Carpenter, Fennema, Ansell & Behrend, 1998). CGI researchers identified changes in the instructional practices of 18 of 21 teachers in the study, and these changes were related to improvements in student achievement. The Elementary Leaders in Mathematics (ELM) project was also able to connect professional development to instructional change and improvements in students' understanding of mathematics (Simon & Shifter, 1991; Shifter & Simon, 1992). Changes in teachers' instructional practices generated changes in the ways students worked and interacted in mathematics class – students became more willing to share strategies, to listen to the strategies of others, and to rely on themselves and their peers as sources of mathematical authority in the classroom. Similarly, following their participation in the Second-Grade Mathematics Project, teachers were more likely to base their instructional decisions on observations of their own students' thinking (Cobb, Yackel, & Wood, 1991; Wood, 1995). Classroom observations indicated that teachers' instructional practices evolved from a directive style of teaching by telling, to an approach of trying not to tell, to a focus on creating opportunities for their students to think and talk about mathematics.

In summary, CGI, ELM, and the Second-Grade Mathematics Project provide existence proofs that professional development experiences can bring about changes in teachers' instructional practices (Simon & Shifter, 1991). Based on their success, this study anticipates that professional development for teachers of mathematics focused on identifying high-level instructional tasks *and* on implementing these tasks in ways that maintain the high-level cognitive demands will lead to both improved teaching practice and improved student learning

outcomes. The following section describes the purpose of the current investigation in greater detail.

1.3. Purpose

According to Henningsen & Stein (1997), learning experiences for mathematics teachers should explicate that task demands can be altered during instruction, and that certain instructional factors serve to maintain or to reduce students' opportunities to engage in high-level cognitive processes. This study will provide exactly that type of learning experiences for teachers of mathematics in hopes of changing teachers' knowledge and instructional practices with respect to the selection and implementation of instructional tasks. The purpose of this study is to determine how professional development focused on the cognitive demands of mathematical tasks will influence mathematics teachers' selection and implementation of the mathematical tasks they use for instruction in their own classrooms. The study will analyze the extent to which the professional development experiences for secondary mathematics teachers enables them to (1) identify the cognitive demands of mathematical tasks; (2) select high-level mathematical tasks for instruction in their own classrooms; and (3) maintain high-level cognitive demands throughout an instructional episode. The hypothesis is that professional development experiences focused on the cognitive demands, selection, and implementation of mathematical tasks will allow teachers to generalize features of instruction that support students' learning of mathematics and can be applied to the selection and implementation of instructional tasks in the teachers' own classrooms.

1.4. Research Questions

This study will analyze changes in teachers' knowledge and instructional practices with respect to the selection and implementation of cognitively challenging mathematical tasks over the course of their participation in a year-long professional development workshop focused on the cognitive demands of mathematical tasks and the implementation of high-level mathematical task during instruction. The study aims to examine the following research questions:

Question 1: Can teachers identify and characterize mathematical tasks as having high-level or low-level cognitive demands, and does this change after participation in professional development specifically focused on the selection and implementation of cognitively challenging mathematical tasks?

Question 2: Do teachers use mathematical tasks with high-level cognitive demands to engage students in learning mathematics, and does this change during and following participation in professional development specifically focused on the selection and implementation of cognitively challenging mathematical tasks?

Question 3: Do teachers implement mathematical instructional tasks in ways that support students' engagement with high-level cognitive demands, and does this change during and following participation in professional development specifically focused on the selection and implementation of cognitively challenging mathematical tasks?

Question 4: What changes (or lack thereof) in teachers' knowledge, selection, or implementation of cognitively challenging tasks can be reasonably associated with individual or group experiences in the professional development sessions?

1.5. Significance

One path toward improving students' mathematical abilities is to improve the quality of the tasks with which students engage during mathematics instruction. If the goal for student learning in mathematics is to reflect the ideals of the NCTM standards (1989, 2000), then students must be exposed to mathematical tasks that provide opportunities to think, reason, and solve genuine problems. Consequently, with the teacher as the agent who chooses the tasks that students encounter in the classroom, this approach depends upon the teachers' ability to select mathematical tasks that encourage thinking, reasoning, and problem-solving. In this study, teachers will be exposed to a framework for analyzing the cognitive demands of mathematical tasks throughout an instructional episode. Increasing teachers' knowledge of how tasks evolve during a lesson and how tasks influence students' learning is expected to generate changes in the way teachers select and implement instructional tasks in their own classrooms, thereby increasing students' opportunities to learn mathematics with understanding. This study will provide empirical data, based on artifacts and observations from teachers' classrooms, about what teachers learn from professional development, whether that learning leads to changes in their practice, and how their learning and instructional change relate to the professional development experiences.

This study will add to the current knowledge base on the effectiveness of using an underlying theoretical framework as the basis for professional development, as called for by Ball

& Cohen (1999) and providing this framework to teachers to serve as a catalyst for teacher learning, reflection, and change. This study extends the work of QUASAR by employing a framework created for research (i.e., Stein, Grover, & Henningsen; 1996; Henningsen & Stein, 1997) to serve as a tool for professional development and assessing whether learning opportunities for mathematics teachers focused on the cognitive demands of mathematical tasks enables teachers to select and implement high-level tasks in their own classrooms,

This study also builds on recent research by Arbaugh (2000; Arbaugh & Brown, 2002), in which the notion of the cognitive demands of mathematical tasks was used as the basis for promoting learning and change in teacher study groups. Arbaugh's research reports changes in percentages of high-level tasks used for instruction in the teachers' classrooms, as will this study; however, in addition to changes in percentages of high-level tasks used in teachers' classrooms, the present investigation will also analyze the teachers' implementation of tasks during instructional episodes by including classroom observations and collections of student work as sources of data.

The work of Matsumura and colleagues (Matsumura, 2003; Matsumura, Garnier, Pascal, & Valdes, 2002; Clare & Aschbacher, 2001; Clare, 2000), in using students work as an indicator of classroom practice will also be drawn upon and extended from reading comprehension to apply to secondary mathematics. In doing so, this study will also provide further data on the effectiveness of the Instructional Quality Assessment (IQA) Toolkit for evaluating academic rigor in mathematics instruction.

1.6. Limitations

This study will examine the effects of a professional development experience on teachers' selection and use of high-level mathematical tasks in their classrooms. Because of a limited number of observations, collections of student work, and collections of tasks, the study is limited to providing a snapshot of instruction in each of the teachers' classrooms at different points in the school year and therefore may not represent the typical instructional practices of individual teachers or of the group of teachers.

The study will provide pre- and post-data in an attempt to portray the influence of the professional development sessions on teacher's instructional practices, though acknowledging the difficulty in establishing a causal relationship between changes in teachers' practice and their experiences in professional development. External influences, such as their experiences in a concurrent university course or professional development program or changes in the leadership in their school district, cannot be controlled for and thus may be responsible for generating the changes in participants' teaching practice identified by this study.

Finally, the small sample size of this study will limit the generalizability of its findings.

1.7. Overview

This document consists of five chapters. Chapter 1 presents the need for the research being undertaken in this study based on the need to improve students' opportunities to learn mathematics with understanding and of the value of high-level tasks in influencing students learning. Chapter 2 provides a more detailed account of the relevant literature on the frameworks used in the study, and on current knowledge about task selection and implementation, and on

current knowledge about effective professional development for teachers of mathematics. Chapter 3 presents the methodology and includes a description of the context of the study, the subjects in the study, the data sources, and the analysis procedures. The results of the analysis are reported in Chapter 4. Chapter 5 presents the discussion of the findings, conclusions drawn from these findings, and suggestions for future research.

2. CHAPTER 2: REVIEW OF LITERATURE

Improving students' opportunities to learn mathematics as called for by NCTM *Standards* documents (1989, 2000) and commissioned educational reports (e.g., "*Before Its Too Late*" [USDE, 2000]) requires improved opportunities for teachers to enhance their knowledge of mathematics, of students as learners of mathematics, and of effective mathematics pedagogy in ways that will enable them to enact reform-oriented mathematics instruction (Ball & Cohen, 1999; Fennema & Franke, 1992; NCTM, 1991). Hence, professional development opportunities for mathematics teachers must promote transformative learning -- learning that generates changes in teachers' current conceptions of effective mathematics teaching and learning that subsequently generate changes in their instructional practices (Thompson & Zeuli, 1999). A central premise in this study is that tasks mediate teaching and learning in school classrooms <u>and</u> in professional development settings. Tasks set parameters for what content is to be learned and for how the learner is to engage with that content. Doyle's (1988, 1983) argument that the tasks students encounter during mathematics instruction structure their opportunities for learning can be applied, "by theoretical consistency" (Cobb, Yackel, & Wood, 1991, p. 88), to teachers' opportunities to learn from the tasks they encounter during professional development experiences. The goal of the present investigation is to determine whether professional development experiences that focus on the cognitive demands of mathematical tasks and the maintenance of high-level cognitive demands throughout an instructional episode will influence

the ways in which mathematics teachers select and implement the mathematical tasks they use for instruction in their own classrooms.

Chapter 2 provides a review of literature in four areas that have direct relevance to this study: knowledge bases for teachers of mathematics, professional development for teachers of mathematics, mathematical tasks as professional development and research tools, and theories of teacher learning. The review provides a rationale for the design and implementation of the intervention and for the assessment of teachers' learning. The first section of this chapter describes three knowledge bases that have been identified as essential for teacher of mathematics and how the professional development in the current study is designed to enhance teachers' knowledge in each of these areas. In the second section, studies of professional development for teachers of mathematics will be reviewed and evidence of transformative learning on the part of teachers will be assessed. In doing so, the second section will establish what is currently known about transformative professional development for teachers of mathematics and how the current investigation is based upon and serves to extend earlier work. The third section of this chapter reviews literature on mathematical tasks, the influence of tasks on students' learning, and the potential of using task analysis as a tool for professional development and for research on teaching. This review provides an argument for using the research frameworks developed by the QUASAR Project as the foundation for the professional development experiences and for the analysis of the data collected in this study. In the fourth section of Chapter 2, theoretical perspectives on learning from cognitive psychology and social anthropology will be reviewed to describe how transformative learning might be enacted through professional development experiences and how these theories of learning play interdependent roles in teacher's construction of knowledge. Hence, in describing a perspective on how teachers learn, the fourth

section will establish that this study is based on a view of teacher-learning grounded in theory, appropriate for the population, and consistent with the design and conduct of the intervention.

The chapter concludes by describing how the professional development intervention in this study draws on characteristics of transformative professional development and theories of teachers' learning that are likely to promote instructional change and how the design of the study will provide valid indicators of changes in teachers' learning and instructional practices following their participation in the study.

2.1. Essential Knowledge Bases for Teachers of Mathematics

Teachers' knowledge and beliefs about mathematics and about students' learning of mathematics influence their instructional practices (Borko & Putnam, 1995) and impact how a teacher responds in the moment-to-moment decisions made during an instructional episode, the plans a teacher makes for a lesson or unit of instruction, and the norms teachers set for daily classroom routines and patterns of interaction (Schoenfeld, 1998). In order to effect changes in instructional practice, then, professional development for teachers of mathematics should change teachers' knowledge and beliefs about mathematics, effective mathematics instruction, and students' learning of mathematics. Consistent with researchers, theorists, and teacher educators in mathematics teacher development, this study draws on Shulman's (1986) seminal paper on the knowledge needed for effective teaching to identify three components essential to implementing reform-oriented mathematics instruction: knowledge of *mathematics*, knowledge of effective *mathematics pedagogy*, and knowledge of *students as learners of mathematics* (Ball & Cohen, 1999; Fennema & Franke, 1992; NCTM, 1991). In the current study, the goal of influencing teachers' selection and implementation of high-level mathematical tasks to engage students in

learning mathematics in their own classrooms will require enhancing teachers' knowledge and beliefs in each of these areas. This section will begin with a discussion of the importance of teachers' knowledge of mathematics, mathematics pedagogy, and students' as learners of mathematics. The section will conclude by describing how the present study will enhance teachers' knowledge in these areas.

2.1.1. Enhancing Teachers' Knowledge of Mathematics, Mathematics Pedagogy, and Students' Learning

In order to enact reform-oriented mathematics instruction, teachers of mathematics require a deep understanding of the mathematical content they teach (Ball, Lubenski & Mewborn, 2001; Sherin, 2002; USDE, 2000; Borko & Putnam; 1995; NCTM, 1991). Mathematics teachers who were successful as students of mathematics in both high school and college may need to strengthen or relearn areas of mathematics in ways that go beyond memorization or rote applications of procedures. As products of mathematics curricula notorious for being a mile wide and an inch deep (Schmidt, 1996), teachers of mathematics may not have had opportunities to develop a deep understanding of mathematical concepts (Ball, et al., 2001; Cooney, 1999). Secondary mathematics teachers, in both middle and high school, have been found to lack a deep understanding of important mathematical topics, such as rational number (Post, Harel, Behr, & Lesh, 1991), functions (Stein, Baxter, & Leinhardt, 1990), and geometry (Swafford, Jones, & Thorton, 1997).

Research summarized by Romberg & Carpenter (1986) and Cooney (1999) indicate that simply having a greater *amount* of mathematical knowledge does not appear to make a teacher more effective in the classroom. However, the *depth* and *organization* of a teacher's

mathematical knowledge has been shown to influence instructional practices. Analyses of expert teachers' understanding of functions (Lloyd, 1999; Stein, Baxter, & Leinhardt, 1990), subtraction (Leinhardt, 1987) and fractions (Leinhardt & Smith, 1985) indicates that expert teachers' mathematical knowledge is highly connected and organized. Lloyd and Wilson (1998) contend that a deep, integrated knowledge of functional relationships and their graphical representations enabled the teacher in their study to effectively implement a reform-oriented high-school mathematics curriculum. Conversely, Stein, Baxter, and Leinhardt (1990) found that a fragmented, superficial understanding of functions and graphing contributed to instruction based on memorized rules and procedures and an inability to take advantage of opportunities to engage students in making meaningful mathematical connections.

A teacher's mathematical knowledge should consist of deep understandings of the mathematical principles that underlie mathematical procedures and of the webs of ideas connected to particular mathematical topics (Ball, et al., 2001). This "profound understanding of mathematics" – defined by Ma (1999) as a connected, structured and coherent knowledge core of mathematical concepts – is necessary in order to understand different representations of a mathematical concept, different solution strategies for solving a problem, and students' thinking or misconceptions about mathematical concepts and procedures. Hill and Ball (2004; p. 332-33) refer to these aspects of teachers' mathematical knowledge as "specialized knowledge of content,…unique to individuals engaged in teaching children mathematics," and different from "common knowledge of content" that would be utilized by mathematicians or other adults in general. Hill, Ball, and colleagues have identified a positive association between teachers' knowledge of mathematics for teaching and student learning outcomes (Hill, Rowan, & Ball, 2005). Similarly, expert teachers of mathematics as described in research by Leinhardt and

colleagues (Stein, Baxter, & Leinhardt, 1990; Leinhardt, 1987; Leinhardt & Smith, 1985) and by Lloyd (1999; Lloyd & Wilson, 1998) exhibited enhanced specialized knowledge of content. Understanding the concepts that underlie mathematical procedures enabled teachers in these studies to provide their students with explanations or explorations that developed students' understanding of how and why the procedure worked, when the procedure was appropriate to use, and how the procedures connected to other related, symmetrical, or inverse procedures. A deeply organized and connected knowledge of mathematics enabled the teacher to consider effective methods and representations for engaging students with mathematical ideas, to recognize the mathematics in students' alternative strategies or ways of thinking, or to capitalize on opportunities for students to make meaningful mathematical connections (Lloyd, 1999; Stein, Baxter, & Leinhardt, 1990). Hence, teachers' knowledge of mathematics for teaching is tightly linked to their knowledge of effective mathematics pedagogy and of students as learners of mathematics (Ball, et. al, 2001; Sherin, 2002).

Developing or enhancing teachers' knowledge of effective mathematics pedagogy and of students' learning are requirements for enabling teachers to enact reform-oriented mathematics instruction (NCTM, 2000, 1991; Ball & Cohen, 1999; USDE, 2000). Experiences in traditional mathematics classrooms, whether as students of mathematics or as preservice teachers, collectively serve to form teachers' conceptions of effective mathematics pedagogy (Cooney, 1999). In traditional forms of mathematics teaching, attention to students' thinking is overshadowed by attention to students' answers. Correct answers reflect "mastery" in the sense that the student can reproduce the information or procedures previously provided by the teacher, and incorrect answers indicate that the student needs to be told or shown again. In contrast, if students' responses, both correct and incorrect, are considered to result from rational, systematic

thought, then attending to students' thinking could inform and influence instructional practice. As teachers endeavor to understand the thinking behind students' ideas, errors, and strategies, they increase their own understanding of the mathematical topic and of the effectiveness of the pedagogy in the lesson (Sherin, 2002). For example, considering why students make consistent and recurring errors [as in manipulating algebraic symbols (Matz, 1980) or regrouping in subtraction (Brown & Van Lehn, 1982)] can raise teachers' awareness of students' fragile understandings of the mathematical principles that underlie common mathematical procedures. Sherin (2002) asserts that consideration of students' novel or unexpected approaches and strategies prompts teachers to reorganize their own mathematical understandings. In these ways, analysis of student's thinking can generate teacher learning and changes in teacher's instructional practices (Sherin, 2002; Carpenter, Fennema, Peterson, Chiang, & Loef, 1989). Based on the premise that teachers will act in ways that they believe are beneficial to students' learning (Remillard, 1999; Borko & Putnam, 1995; Thompson, 1992), changing teachers' instructional practices requires changing teachers' knowledge and beliefs of effective mathematics pedagogy.

2.1.2. Enhancing Teachers' Knowledge in the Present Investigation

The professional development experiences that constitute the intervention in the present investigation are intended to enhance teachers' content knowledge and pedagogical content knowledge of mathematics. Though we did not endeavor to teach specific mathematical content, the professional development intervention in this study provided opportunities for teachers to increase their knowledge of mathematics for teaching as described by Hill and Ball (Hill & Ball, 2004): "common knowledge of content" (i.e., by solving and discussing mathematical tasks) and "specialized knowledge of content" (i.e., by considering multiple strategies and the mathematical

connections between them, or by considering the value added from different representations of mathematical concepts). We chose to initiate teachers' learning by engaging them in solving cognitively challenging mathematical tasks. This approach is consistent with many professional development projects that will be reviewed later in this chapter (e.g., Farmer, Gerretston, & Lassak, 2003; Borasi, Fonzi, Smith, & Rose, 1999; Simon & Shifter, 1991). From solving challenging mathematical tasks, teachers have opportunities to deepen their mathematical knowledge and gain first-hand experience as "learners" in reform-oriented mathematics lessons. By reflecting on how the learning experience contributed to their own learning, teachers are able to appreciate the power of cognitively challenging tasks, of persisting in the struggle to make sense of mathematical ideas and solve mathematical problems, and of reform-oriented mathematics pedagogy in supporting their students learning.

In addition to opportunities to solve mathematical tasks, at the heart of the intervention in this study are on-going opportunities for project teachers to engage in assessing the cognitive demands of mathematical tasks and the implementation of mathematical tasks during instructional episodes. Experiences in evaluating the cognitive demands of tasks (i.e., what cognitive processes the student can engage with as they solve the task), the specific connections between mathematical ideas or representations that a task has the potential to elicit, evidence of students making these connections during an instructional episode or in samples of students' work, and the instructional strategies that influenced students' opportunities to make these connections collectively serve to develop teachers' pedagogical content knowledge (Ball, et al., 2001; Sherin, 2002; Shulman, 1986). Teachers' knowledge of the cognitive demands of mathematical tasks and teachers' decisions in selecting instructional tasks are aspects of pedagogical content knowledge referred to as "knowledge of content and teaching" (Hill,

Rowan, & Ball, 2005; Hill & Ball, 2004); considering what questions to ask to assess and advance student's thinking, or considering how to select and order presentation of students' work are aspects of "knowledge of students and content."

In the present investigation, providing teachers with opportunities to solve challenging mathematical tasks and to evaluate the cognitive demands of mathematical tasks (as printed and as implemented during an instructional episode) is intended to impact teachers' ideas of how students learn mathematics and, in turn, of how they should endeavor to teach mathematics. A desired outcome is that teachers will begin to base their notions of effective mathematics teaching upon pedagogical moves that support the maintenance of high-level cognitive demands throughout an instructional episode (i.e., the classroom-based factors identified by Stein, Grover, & Henningsen, 1996). Focusing on the cognitive demands of mathematical tasks also closely connects mathematics and students' thinking as teachers engage in differentiating tasks based on opportunities for students' learning, assessing the cognitive processes students actually engaged in while solving the task, and determining the mathematical understanding present in samples of students' work. Hence, engaging teachers as learners in the types of mathematics lessons that they are intended to implement in their own classrooms and in analyzing the cognitive demands of mathematical tasks (in print and as the tasks play out during instruction) promotes teachers' learning in ways that connect teachers' knowledge of mathematics, mathematics pedagogy, and students as learners of mathematics.

Specific details on the professional development experiences provided to teachers participating in this investigation will be provided in Chapter 3. In the next section, the approach to enhancing and analyzing teachers' knowledge and instructional practices undertaken by prior

professional development studies will be reviewed in order to inform the design of the intervention and methodology of the present investigation.

2.2. Review of Professional Development Projects that Inform this Investigation

In order to impact teachers' instructional practices, professional development projects often intend to enhance teachers' knowledge in all three of the essential knowledge bases for teachers of mathematics identified in the previous section (i.e., mathematics, mathematics pedagogy, and students as learners of mathematics); though different projects choose different entry points to initiate the conversation with teachers (Borasi, et al., 1999). For example, a project might begin by challenging teachers' conceptions of mathematics (e.g., Swafford, et al., 1997; Simon & Shifter, 1991), by introducing teachers to students' ways of thinking and learning (e.g., Cobb, Yackel, & Wood, 1991; Carpenter, et al., 1989), by engaging teachers in analyzing inquiry-based, reform-oriented mathematics instruction (e.g., Wallen & Williams, 2000; Barnett, 1998), or by assisting teachers' implementation of reform-oriented mathematics pedagogy and/or curricula (e.g., Sherin, 2002; Smith, 2000; Remillard, 1999). The professional development projects that will be discussed in this section used cognitively challenging tasks, narrative cases of mathematics instruction, reform-oriented mathematics curricular materials, and/or examples of students' mathematical thinking as tools for strengthening teachers' knowledge of mathematics, for changing teachers' notions of how mathematics is best taught and learned, and for catalyzing changes in teachers' instructional practices.

This section will review professional development studies for teachers of mathematics and assess the evidence of enhancing teachers' knowledge and implementation of effective mathematics pedagogy. As an organizing structure, the studies are presented along the

continuum of (1) studies that fostered changes in teachers' knowledge and beliefs but did not assess teachers' instructional practices, (2) studies that utilized teachers' self-reports as indications of instructional change, and (3) studies that utilized classroom artifacts and observational data as evidence of instructional change. Hence, this section describes prior professional development research and explicates how this study utilizes and extends earlier efforts at fostering and analyzing teacher learning and instructional change.

2.2.1. Changing Teachers' Knowledge and Beliefs

The professional development studies reviewed in this section enhanced teachers' knowledge and beliefs by engaging teachers' in the analysis and discussion of classroom episodes, or instructional cases, of mathematics teaching and learning. An *instructional case* is a narrative or video depiction of a teaching episode or event created for use in professional development settings (Merseth, 1996). According to Merseth (1996), instructional cases are created with the explicit intention of stimulating thought and debate, and should contain enough detail and decision points to make for interesting discussions and arguments. The studies reviewed in this section used case discussions to foster changes in teachers' knowledge and beliefs that were intended to generate changes in classroom practice. For example, based on experiences in the Mathematics Case Methods Project (MCMP), Barnett (1998; 1991) concludes that case discussions create a climate that is conducive to informed strategic inquiry, an investigative process where teachers endeavor to understand and resolve issues for themselves, grounded in their understanding of mathematics and how mathematics should be taught and learned. The goal of the MCMP is to develop teachers' ability to draw on their understanding of mathematics and of students as learners of mathematics to inform instructional decisions. In

particular, Barnett analyzed 27 case discussion transcripts with elementary and middle school teachers in order to determine common thematic dimensions that arise within the case discussions. Barnett concluded that teachers gained an appreciation of the difficulties students experience with fractions, began to focus on developing students' learning of mathematics, and began to consider how to sequence instructional tasks in order to develop students' understandings. Barnett attributes the changes in teachers' knowledge of mathematics pedagogy to the opportunity to engage in pedagogical reasoning provided by the analysis and discussion of the case and the "collaborative construction process" of participating in group deliberations where teachers consider ideas that had not occurred to them as individuals.

The work of Wallen and Williams (2000) also indicates that cases can foster the disposition for a stance of inquiry toward ones' own teaching. In their analysis of the notes and comments (recorded by project staff) produced during case discussions by 115 teachers implementing a reform-oriented, integrated curricula in 9th-12th grade mathematics, Wallen and Williams found that the teachers' reflections on the mathematics pedagogy in cases often became personalized (i.e., as indicated by the use of "I" or "my" rather a pronoun appropriate to the teacher in the case). Through the case discussions, participants were able to identify and attempt to solve problems they had been wrestling with in their own classrooms. Taken together, the results reported by Barnett (1998) and Wallen and Williams (2000) indicate that case discussions provide opportunities for teachers to examine mathematics pedagogy that challenges traditional views of effective teaching and learning, without having to focus on their own practice. In this way, the intervention in these studies affected change in teachers' knowledge and beliefs. These studies did not follow teachers into their classroom subsequent to the case experiences and thus do not make claims of changing teachers' instructional practices.

This study builds on the findings from Barnett (1998; 1991) and Wallen and Williams (2000) of the value of using narrative cases to foster mathematics teachers' reflection on reform-oriented instructional practices.. The goals for teachers' learning from the case analysis and discussions are consistent with the overall goals of this study – enhancing teachers' knowledge, selection, and implementation of cognitively challenging tasks.

2.2.2. Evidence of Change in Teachers' Knowledge and Instructional Practices based on Teachers' Self-Reports and Informal Classroom Observations

The professional development studies reviewed in this section provide evidence of changes in teachers' knowledge and beliefs and use teachers' self-reports and informal classroom observations (i.e., lessons were observed but not analyzed as data) as evidence of changes in teachers' instructional practices. Several of these projects engaged teachers as learners in reform-oriented mathematics lessons to initiate changes in teachers' views of effective mathematics teaching and learning. In Project LINCS (Swafford, et al., 1997), for example, middle school mathematics teachers participated in a content course in which they solved cognitively challenging geometry tasks. Project teachers also participated in a research seminar focusing on studies of students' cognition and knowledge of geometry (i.e., van Hiele levels of geometric thinking). Results from pre- to post-test of the teachers' depth of geometric knowledge indicated that 72% of teachers gained at least one van Hiele level, and 56% of the teachers gained two levels. In another pre- and post-assessment, teachers were provided with a two-page lesson copied from the teachers' edition of a geometry textbook and given 20 minutes to write a lesson plan, state their goals for the lesson and their expectations for students, and to indicate how they would alter the lesson from the suggestions provided by the textbook. While the specific content

of the lesson was embedded in the geometry course, the teaching of this content was not explicitly modeled or addressed within the professional development sessions. The pre- and post-lesson plans were analyzed with respect to (1) the van Hiele level of the lesson tasks and of the teachers' goals and expectations for students and (2) the presence of reform-oriented pedagogy in the teachers' lesson activities and alterations to the text lesson. Comparisons of the tasks in the pre- and post- lesson plans indicated a significant decrease in tasks at van Hiele Level 1 and a significant increase in tasks at van Heile Level 2. Though the tasks overall increased by one level, they remained at a low level of cognitive demand as classified on the van Hiele taxonomy. Tasks with the potential to elicit higher levels of cognitive demand (i.e., van Hiele levels 3, 4, and 5) were equally absent in both pre- and post-lesson plans, with a maximum of 3% of tasks at Level 3 and no tasks at Levels 4 and 5. In the post- lesson plans, however, there were more changes to the printed lesson in the text -- teachers appeared more confident in making substantial deletions and insertions, the nature of which tended to incorporate greater student exploration and use of manipulatives; though, again, these changes typically did not increase the cognitive demand of instructional tasks beyond van Hiele Level 2. Overall, Project LINCS increased teachers' knowledge of mathematics by 1 to 2 van Hiele levels and enhanced teachers' knowledge of mathematics pedagogy and students' thinking in geometry in the sense that, following their participation in the project, teachers were able to plan lessons and to modify text lessons in ways that incorporated a higher level of geometric thinking and more aspects of reform-oriented mathematics pedagogy than was evident in their lesson plans prior to participation in the study. In this way, Project LINCS was successful in changing teachers' image of quality mathematics instruction in geometry (Schoenfeld, 1998).

Did involvement with Project LINCS also change teachers' instructional practices? Observations were conducted in project teachers' classrooms at 3-5 points during the subsequent school year. The researchers state that, "based on teacher reports and on teacher and researcher perceptions" (p. 476), project teachers were (1) spending more time and more quality time on geometry instruction; (2) more willing to try new ideas and instructional approaches; (3) more confident to respond to higher levels of geometric thinking; and (4) more likely to engage in risk taking that enhanced student learning. Case studies of four teachers provide examples of changes in the teachers' instructional practices that can be reasonably traced back to the teachers' experiences in the project. For example, one teacher incorporated a task on tessellations, something she had not previously taught but had experienced as a learner in the LINCS geometry course. Another teacher talked explicitly about the van Hiele levels of the tasks she used during the observed lesson and about how she made task adaptations based on her new knowledge of the van Hiele levels of geometric thinking. Teachers' self-reports provide evidence of changes in knowledge and beliefs that the teachers and researchers both attribute to the project's intervention. The researchers assert that "teachers were incorporating new geometric ideas and tasks into their programs …consistent with our observations of both their instruction and their assessment tasks" (p. 477); however, very limited observational data is provided to substantiate these claims. Results reported in the study offer no assessment of the van Hiele levels of the instructional tasks used in the observed lessons nor of the level of implementation of those tasks during instruction (i.e., the van Hiele level of students' actual thinking as they engaged in solving the tasks). Overall, the results provided by Swafford and colleagues make it difficult to ascertain whether teachers in the LINCS project changed their instructional practices toward the intended goals of the project (i.e., whether the tasks and lessons used during geometry instruction

increased in van Hiele level of cognitive demand and in aspects of reform pedagogy). To substantiate claims of instructional change, the observed lessons might have been analyzed for the van Hiele level of the tasks, the van Hiele level of the implementation of the tasks, and the aspects of reform-oriented mathematics pedagogy present during the lesson.

Similarly, the SummerMath for Teachers (Schifter & Simon, 1992) and the Educational Leaders in Mathematics Project (ELM) (Simon & Shifter, 1991) provide evidence of enhancing teachers' knowledge and beliefs, but do not utilize classroom artifacts or observations to inform their analysis of changes in teachers' instructional practices. Teachers in both projects were provided with opportunities to examine the nature of mathematics and the process of learning mathematics as a basis for developing new ideas about effective mathematics teaching and learning. Teachers attended a two-week summer institute in which they participated in solving challenging mathematical problems (e.g., high school mathematics teachers engaged with the content of weighted averages and direct/inverse variation). Each "lesson" was followed by an explicit discussion of roles of the teacher (i.e., the professional development facilitator), the students (i.e., the project teachers), the structure of lesson, and how all of these facets impacted the learning experience. Teachers also engaged in designing lesson sequences intended to provide opportunities for students to construct mathematical ideas. During the school year following teachers' participation in the SummerMath and ELM projects, project staff observed one lesson per week in the teachers' classrooms and conducted post-observation interviews.

Teacher writings and interviews following their participation in the ELM intervention provided strong indications that solving challenging mathematical tasks and being involved in constructivist learning experiences stimulated changes in the teachers' personal views of themselves as mathematics learners, about how mathematics is learned, about how mathematics

should be taught. Based on the common themes in their writings, teachers who participated in the ELM study claimed to have developed a more critical perspective on their own practice, changed their views of themselves as mathematics learners, implemented new instructional strategies, and developed new views of students' learning of mathematics. Some teachers' writings provided evidence of reflection and meta-analysis of the development of their own learning during the ELM sessions and how these new insights influenced their ideas about their students' learning. Teacher interviews indicated that, during the school year following the initial ELM intervention, 92% of project teachers were implementing strategies modeled in ELM, and 52% of the teachers continued to implemented these strategies consistently 2 years later. Furthermore, based on the interviews, 64% of teachers were considered to base their instructional decisions on a constructivist view of learning. Overall, the evidence from the teachers' self-reports (i.e., their writings and interviews) indicates that teachers participating in the ELM Project began to consider their changing role as teachers in the classroom in light of their new experiences as learners in the professional development institute -- teachers implemented new instructional strategies modeled in the ELM intervention and were developing a constructivist epistemology.

However, the SummerMath and ELM Projects did not analyze the classroom observations as a source of data. Rather, only teachers' writings and post-observation interviews were examined to study the impact of the intervention on project teachers' "learnings, understanding, and implementation" (Simon & Schifter, 1991, p. 317). Teachers' writings were analyzed to identify data relevant to the project's impact on teachers, and this data was categorized thematically. Teacher interviews were evaluated to assess teachers' implementation of strategies acquired from the ELM intervention and teachers' level of constructivist epistemology. Observational data is not identified as a source of evidence in the discussion of

teachers' instructional change. In this sense, the SummerMath, ELM, and LINCS projects (Swafford, et al., 1997) assessed *implementation* without analyzing classroom artifacts or observations to substantiate teachers' self-reports. Extending the research design to include an analysis of classroom artifacts and observations would have provided further evidence of the effectiveness of the type of intervention used in these studies -- professional development grounded in a solid conceptual framework of teachers' learning, students' learning, and effective mathematics pedagogy. The current investigation builds on the strengths of the conceptual underpinnings of prior research and extends the scope of analysis by utilizing classroom data to provide evidence of teachers' learning and instructional change following their participation in this type of professional development.

Professional development studies conducted by Farmer, Gerretston, and Lassak (2003) and Borasi, Fonzi, Smith, and Rose (1999) also observed lessons in the project teachers' classrooms but did not utilize this data in their reports of teacher learning and instructional change. In the Enhancing Mathematics in the Elementary School (EMES) Project (Farmer, et al., 2003), teachers engaged as learners in mathematics lessons resonant of the type of mathematics instruction that was intended for teachers to implement in their own classrooms. Teachers were provided with opportunities to reflect on their own learning experiences and to make explicit connections between their experiences as learners in the professional development and students as learners in their classrooms. Farmer and colleagues (2003) describe teachers' learning from professional development in terms of three "levels of appropriation." At Level 1, teachers appropriate specific content (instructional units, tasks, activities) and pedagogical techniques (i.e., have students present solutions) that they will implement "as is," while other aspects of their teaching and their ideas of effective mathematics pedagogy remain unaffected. Teachers at

level 1 gain tools to add to repertoire that do not generalize beyond the specific ways in which the tools were originally encountered. At Level 2, teachers appropriate general "lessons learned" that more broadly influence their instructional practices. Level 3 appropriation constitute an inquiry approach to instruction in which teachers not only generalize and implement lessons learned from the professional development experience, but also persist in refining the implementation of these experiences in their own classroom. They continuously reflect on and learn from their own teaching practice.

Though Farmer and colleagues (2003) observed at least 2 lessons in each teacher's classroom, teachers' reflections and comments are used to assess teachers' levels of appropriation. Self-reports in one case study indicate that the teacher had begun to implement explorations and other standards-based teaching techniques that she personally attributed to her experience in the EMES project (i.e., the teacher used real-world examples, technology, grouping, manipulatives, etc.). The second case study is largely based on the reflections and writings of the individual teacher, though the researchers indicate that following participation in the EMES project, classroom observations indicated that the teacher assessed and modified her practice based on student thinking and personal reflections. However, it is unclear the extent to which data from classroom observations were utilized in forming this conclusion. In contrast, Farmer and colleagues utilize classroom observations in the third case study to provide an account of a veteran 4th grade teacher implementing a cognitively challenging instructional task. The teacher's reflections on her own learning during the EMES project, on her students' learning of mathematics, and on the connections between them provide evidence that she had appropriated from the project exactly the goals that the project developers had intended. Consistent with earlier case studies, no details are provided to substantiate the researchers'

claims of instructional change (i.e., rich mathematical discourse among students and between student and teacher) and no pre-data (other than the teacher's reflective self-reports) is provided to confirm that the instructional practices resonant of reform-oriented pedagogy are indeed changes in the teachers' prior methods of instruction.

In the professional development study by Borasi, Fonzi, Smith, and Rose (1999), the intervention was designed to encourage middle school mathematics teachers to implement reform-oriented, inquiry-based mathematics instruction for all students in their classrooms, including mainstreamed learning-disabled students. Teachers in the project participated in a summer institute in which they engaged (first as learners) with three inquiry-based units of mathematics instruction and then received support in implementing these and other inquiry-based units into their own classrooms throughout the following school year. Analyses of teachers' written artifacts and self-reports indicated that project teachers valued and practiced inquiry-based instruction; and classroom artifacts are utilized to assess implementation and to supplement self-reports. For example, results indicate that all 39 teachers implemented at least one inquiry-based unit, with 18 teachers implementing 2-3 units and 13 teachers implementing more than three units. Inserting the new units into their current curricula served to increase the percentage of time that instruction was based on high-level mathematical tasks (ranging from 1.25% to 60.0% of instructional time, with an average of 22%) and provided students in the project teachers' classrooms with the exposure to high-level tasks that is a crucial first step in promoting students' learning of mathematics with understanding (Stein & Lane, 1996). Additionally, 24 of the 39 teachers designed their own inquiry units, and 16 devoted more than 25% of instructional time (i.e., 10 weeks or more) to inquiry-based instruction. This would imply that many of the project teachers were able to generalize the principles and ideas embedded in

the specific instructional materials to apply to their curriculum and instructional practices more broadly. A necessary second step in fostering students' learning of mathematics is the implementation of cognitively challenging tasks in ways that maintain students' opportunities to engage in high-level cognitive processes (Stein & Lane, 1996; USDE-NCES, 2003), and Borasi and colleagues acknowledge that the inquiry units were not consistently implemented as intended. According to the researchers, visits to the teachers' classrooms "suggest that these experiences represented a substantial step forward toward implementing the vision for school mathematics articulated in the NCTM Standards (1989), although they could not all be considered legitimate examples of inquiry teaching (1999, p. 63)." Hence, the study provides evidence that teachers' increased their use of inquiry-based units, but did not analyze the implementation of the inquiry units during instruction in the teachers' classrooms.

Comparisons of the studies presented above, as well as consideration of each study's strengths and weaknesses, serve to inform the present investigation in several ways. First, the discussion of teachers' levels of appropriation by Farmer and colleagues (2003) provides a framework for explaining how different teachers participating in the same professional development experience will take away different learning experiences. Similarly, Simon and Shifter (1991) identified ELM teachers whose learning consisted of adding teaching strategies and materials to their repertoire and those who had fundamental shifts in their view of mathematics teaching and learning, reflecting Thompson and Zeuli's (1999) descriptions of additive and transformative learning, respectively. The levels of appropriation highlight the importance of considering teachers as learners in professional development settings – teachers' disposition or orientation toward their own learning impacts what they will appropriate from the professional development experience. This has important implications for the design of

professional development experiences that are intended to be transformative – teachers should engage in activities that promote generalizations from specific professional development activities that apply to teaching more broadly, and teachers should be provided with opportunities to reflect on their own teaching and their own learning. Second, a framework for analyzing teachers' learning and instructional change consistent with the goals of the professional development project facilitates connections between changes in teachers' knowledge, beliefs, and instructional practice and their experiences in the professional development intervention. For example, Simon and Schifter (1991) assessed the level of use of instructional strategies modeled in the ELM project and the level of constructivist epistemology, and Swafford and colleagues (1997) assessed the van Hiele level of the teachers' lesson plan. Third, analyzing classroom artifacts and observations consistent with the goals of the study would have enhanced the design and the results of the studies identified in this section. For example, Swafford and colleagues could provide the van Hiele levels of the tasks and the task-implementation in the lessons observed in the project teachers classrooms, and Simon and Schifter could analyze classroom observations using their scales for the presence of a constructivist epistemology and for the levels of use of instructional strategies. In this way, self-reports and informal observations could be substantiated by classroom artifacts and observations that were systematically collected and analyzed, and the researchers could make a claim about the effectiveness of their analytical tools as valid indicators of classroom practice. Collectively, the studies raise the question of whether self-reports constitute evidence of instructional change. Borasi and colleagues (1999) pose the question, "What were the effects of the program in terms of changes in participants' beliefs and practices? The answer to this question may be considered the ultimate measure of success for teacher development programs" (p. 58). The studies

reviewed in this section provide evidence of important changes in teachers' knowledge and beliefs based on teachers' self reports (i.e., interviews, writings, surveys, and in some cases, pre/post data), changes that the teachers' themselves frequently attribute to their participation in the respective projects. However, several of the studies also base their evidence of changes in teachers' instructional practices on self-reports rather than on actual classroom data, and did not assess teachers' implementation of reform-oriented instructional materials or strategies.

The present investigation will collect pre- and post-data of teachers' knowledge and instructional practices with respect to the specific goal of the study – to influence teachers' selection and implementation of cognitively challenging mathematical tasks for instruction in their own classrooms. Classroom artifacts (i.e., instructional tasks and student work) and observations will be the centerpiece of the analysis in this study, supplemented by artifacts from the professional development sessions and teachers' self-reports. In this way, the design of the present investigation contrasts the studies reviewed in this section by triangulating observational data and classroom artifacts with teachers' self-reports. Additionally, the level of detail in the analysis of observational data and reports of the findings will assess evidence of teachers' learning and instructional change specific to the intended goals of the project – the level of cognitive demand in the tasks teachers select and implement for instruction in their own classrooms.

2.2.2.1. Evidence of Changes in Instructional Practices Based on Classroom Observations

The professional development studies reviewed in this section provide evidence of changes in teachers' instructional practices based on classroom observations. In the professional development study conducted by the Cognitively Guided Instruction project (CGI) (Carpenter, et

al., 1989), researchers observed a minimum of 16 lessons in each teacher's classroom during the school-year following teachers' participation in the CGI summer workshop. Forty 1st-grade teachers participated in the study -- 20 attended a 4-week summer workshop intended to familiarize them with findings of research on young children's development of addition and subtraction strategies and to provide opportunities to plan instruction based on this knowledge; the other 20 served as the control group and attended two 2-hour workshops focused on problem solving. All 40 teachers were observed during the following school year. The goal of the study was to determine the impact on teachers' beliefs and instructional practices of professional development that exposed them to a research framework for understanding and analyzing children's mathematical thinking that could also form the basis for instructional decisions.

In addition to assessing teachers' instructional practices, the researchers also collected classroom artifacts and data on teachers' knowledge and beliefs. An outstanding feature of the analytical tools used by CGI researchers is that they specifically address the goals of the study. For example, teacher's knowledge of students' thinking was measured by asking teachers to predict how individual students in their own classrooms would solve specific problems and whether the student would obtain the correct answer, and teachers' predictions were then compared to the students' actual responses. Changes in teachers' beliefs were measured by a pre- and post-questionnaire, administered to both the CGI and the control group, designed to assess their assumptions about teaching and learning addition and subtraction. Teachers were also asked to plan a unit of study and to create a year-long plan for instruction in addition and subtraction based on CGI principles. Classroom observations were conducted in 4 separate week-long observation periods between November and April.

Results on teachers' knowledge of students thinking show that CGI teachers were better able to predict their students' strategies, indicating increased attention to students' development of mathematical ideas than teachers in the control group. In fact, teachers in the control group over-predicted students' use of memorized number-fact strategies by two to three times and consistently predicted that students would have a much higher recall of number facts than was indicated by the students' actual strategies. This implies that teachers in the control group were not aware of the thinking of students in their classrooms. Similarly, scores on the belief scales from the pre- and post-questionnaires indicated that the CGI teachers had become more cognitively guided in their beliefs about children's learning than their peers in the control group than they had been prior to the CGI workshop, though both groups were rated as becoming more constructivist in their beliefs from pre to post. Classroom observation results indicate that teachers who participated in the CGI workshop based instruction in addition and subtraction on word problems significantly more frequently than the control group (54.58% vs. 36.19%) and significantly less frequently on number fact problems (25.95% vs. 47.20%). Furthermore, CGI teachers (1) spent significantly more time problem-posing and listening to students' explanations, and thus spent significantly less time providing feedback on answers, and (2) more frequently allowed students choice of strategy, thus significantly less frequently directed students toward the use of advanced counting strategies. Overall, the findings from CGI provide evidence that exposing teachers to research on children's thinking influenced the teachers' knowledge of children's development of addition and subtraction strategies, their beliefs about teaching and learning of addition and subtraction, and their instructional practices in ways consistent with CGI goals of the researchers (i.e., teachers attended to and based instructional decisions on students' thinking as modeled in the CGI summer workshop).

The QUASAR Project (Silver & Stein, 1996) analyzed teacher learning and instructional change through extensive observations and documentation of middle school mathematics teachers' efforts to implement reform-oriented mathematics instruction (Silver & Stein, 1996; Smith, 2000; Stein, Grover, & Henningsen, 1996). The goal of QUASAR was to reform mathematics instruction in diverse, economically challenged urban areas in ways that provided students with opportunities to think, reason, and problem-solve. Project teachers worked together with administrators and university resource partners to "develop, implement, and refine innovative mathematics instructional programs for all students" (Smith, 2000, p. 354). Professional development for teachers participating in QUASAR consisted of coursework, workshops, professional meetings, collaborative activities with colleagues, classroom-based support, and individual reflective activities (Brown, Smith, & Stein, 1996). These professional development activities were designed to support teachers' comprehension (i.e., knowledge and beliefs) and their instructional practices (transformation, implementation, and reflection) (Brown, Smith, & Stein, 1996) as they endeavored to implement an innovative mathematics curriculum in ways that would provide students with opportunities for thinking, reasoning, problem-solving, and communication. The focus of the professional development activities was to allow teachers to engage with the curriculum as learners and to refine their own implementation of the curriculum by analyzing and reflecting upon samples of students' work and videotaped instructional episodes.

In their analysis of the cognitive demands of the tasks teachers used for instruction and of the ways in which these tasks were enacted by teachers and students during instruction, Stein, Grover, and Henningsen (1996) concluded that QUASAR teachers were successful in selecting and setting up cognitively challenging tasks for their students. Analysis of the cognitive demands

of a random sample of 144 tasks across QUASAR sites indicated that 74% of the tasks teachers selected for instruction had high-level cognitive demands. QUASAR teachers had more limited success in maintaining high-level cognitive demands during implementation, with 42% of cognitively challenging tasks enacted in ways that maintained students' opportunities to engage with high-level cognitive processes.

Several research studies provide evidence of the impact of the QUASAR intervention in effecting changes in teachers' knowledge and instructional practices. Research by Smith and colleagues (Smith, 2000; Stein, Smith, & Silver, 1999) presents a detailed analysis of the learning of one middle-school teacher participating in the QUASAR Project. The teacher attended workshops to support the implementation of a reform-oriented, conceptually-based mathematics curriculum. The teacher generalized the pedagogical approach to apply to mathematics instruction broadly, and even recognized similarities to her long-established ways of teaching language arts (Smith, 2000). However, she initially doubted the new approach and its benefit to students' learning of mathematics. Rather than implementing the tasks as intended by the reform-oriented curriculum, she simplified the challenging aspects of the tasks to a focus on following procedures. Through the school year, the teacher engaged in professional development experiences consisting of opportunities to reflect on instructional practice, analyze students' work, and solve challenging mathematical tasks. As the teacher reflected on her teaching and on the reactions of her colleagues to videotaped segments of her lessons, she began to see the need to change her questioning approach (i.e., to move away from directive questions and choral response), to give students longer periods of time to engage in solving problems, and to let students question each other to clarify their misunderstandings. Smith (2000) identifies subsequent changes over the course of the school year in the teachers' directiveness, in student

participation and engagement in the lessons, in the teachers' definition of success, and in students' opportunities for problem-solving. Opportunities to reflect on and learn from her own teaching provided by the QUASAR intervention (i.e., the meetings and collaborative interactions with her peers and university resource partners) catalyzed changes in the teachers' instructional practices.

Brown, Smith, and Stein (1996) compare the nature and extent of professional development support with actual changes in instruction in project teachers' classrooms with respect to the goal of implementing reform-oriented tasks and instruction. Teachers in three QUASAR sites (Sites A, B, and C) had extensive and ongoing experiences to develop their comprehension of mathematics, pedagogy, and student thinking, but such support was only minimally available to teachers at Site D. Additionally, teachers in Site A use videotaped lessons and samples of students' work to encourage discussions amongst colleagues and encourage reflection. Classroom observations in each of four project sites provide evidence of the extent to which instructional practices were consistent with the goals of the QUASAR study. Analysis of teachers' selection of high-level instructional tasks indicated that instruction in Site A was the most consistently based on cognitively challenging tasks (94%), Site D was the least (50%) and Sites B and C fell almost directly in between (~75%). Implementation of high-level tasks in ways that maintained the cognitive demands provided similar results, with tasks at Site A more consistently maintained (61%) than Sites B and C (43% and 33%, respectively) and Site D the lowest (11%).

The QUASAR results for the selection and implementation of cognitively challenging tasks appear far better than that of recent national studies of the quality of mathematics instruction in U.S. classrooms. In the sample of classrooms analyzed by Horizon Research

(Weiss & Palsey, 2004; Weiss, Pasley, Smith, Banilower, & Heck, 2003), only 15% of observed lessons were classified as providing opportunities for thinking, reasoning and sense-making in mathematics. The percentages for the selection and implementation of cognitively challenging tasks by QUASAR teachers are also high in comparison to TIMSS data (17% selected, <1% implemented faithfully). Possible explanations include different definitions of *task* – TIMSS counted individual problems as individual tasks (USDE-NCES, 2003) where QUASAR considered sets of similar problems as the same task (Stein, Grover, & Henningsen, 1996). Another difference is the presence of support for QUASAR teachers vs. teachers at large in the TIMSS study. Many QUASAR teachers elected to participate in the project; hence, they may have had greater motivation and commitment toward reform-oriented instruction than a random selection of teachers. The higher percentage of cognitively challenging tasks used by QUASAR teachers may be attributed to professional development opportunities that supported teachers' knowledge and beliefs about the value of basing mathematics instruction on high-level tasks. Likewise, higher percentages of implementation that maintained the cognitive demands might be due to professional development that supported changes in teachers' instructional practices (i.e., planning, assessment, implementation, and reflection) and changes in teachers' beliefs about how mathematics should be taught and learned. The results provide evidence of teachers' learning and instructional change following their participation in the QUASAR project, and indicate that these experiences enabled them to enact reform-oriented instruction to a greater degree than if the intervention from QUASAR had not been present.

Several frameworks emerged from the analysis of observational data in the QUASAR project. The following section will describe the QUASAR frameworks and how the current study

uses them as tools for professional development and for the analysis of classroom artifacts and observations.

2.3. QUASAR Frameworks as a Basis for Professional Development and Research

As described above, the QUASAR Project sought to increase the level of mathematical understanding and achievement of students in urban, disadvantaged communities. Accomplishing this feat required improving students' opportunities to learn mathematics, which in turn required improving instructional tasks and the ways in which students engaged with those tasks in the process of learning mathematics (Doyle, 1988, 1983; Stein & Lane, 1996; Hiebert & Wearne, 1993). This section will present the frameworks that were developed by QUASAR researchers to analyze the connection between teaching and learning in QUASAR classrooms. In doing so, this section will describe the influence of mathematical tasks on students' learning of mathematics, the value of using the QUASAR frameworks in professional development with teachers of mathematics, and the validity of using these frameworks to guide the collection and analysis of classroom artifacts and observations.

2.3.1. The QUASAR Frameworks

Researchers involved in the QUASAR Project conducted hundreds of classroom observations throughout the first 5 years of the project (1990-1995). The research reviewed in this section drew from a data base of 324 classroom observations conducted during the first three years of the project (3 sets of 3-day observations in 3 representative teachers' classrooms in each of the 4 initial project sites per year) (Stein, et al., 1996; Stein & Lane, 1996; Henningsen & Stein, 1997). From this data base, a stratified random sample of 144 classroom observations were

selected for analysis, equally distributed across season, teacher, year, and project site and reflecting the percent of lessons at each grade level in the entire data base. Based on these classroom observations, QUASAR researchers analyzed teachers' instructional practices with respect to the selection and implementation of the tasks teachers used for instruction. Stein and colleagues (2000) note that two central premises are necessary for an analysis of instruction based on instructional tasks:

(1) different tasks require different levels and kinds of thinking; and

(2) the cognitive demands of tasks can change throughout an instructional episode (p. 3).

These two premises form the basis of the QUASAR frameworks for analyzing the selection and implementation of mathematical tasks in the classroom observations, respectively, and will be described below.

Analyzing instructional tasks. Different mathematical tasks place different demands on students' thinking. According to Doyle (1988; 1983), a useful framework for describing and analyzing students' academic work is in terms of the cognitive level of instructional tasks, defined as "the cognitive processes students are required to use in accomplishing the task" (1988, p. 170). Content labels are not useful in describing the tasks students are asked to accomplish during instruction; for example, "multiplication" can mean different things under different expectations and given different resources. Instead, Doyle (1983) identifies four categories of instructional tasks (memory tasks, procedural or routine tasks, comprehension/understanding tasks, and opinion tasks), which he organizes into two cognitive levels of academic work. Lower levels of academic work include memory tasks and procedural or routine tasks. These tasks often involve the memorization or application of formulas or algorithms (Doyle, 1988). Comprehension/understanding tasks represent higher cognitive levels of academic work and

engage students with cognitive processes such as comprehension, interpretation, flexible application of knowledge and skills, selection of strategies to solve problems, assembly of information from several sources to accomplish the task, drawing inferences, and formulating and testing conjectures (Doyle, 1988).

Drawing on Doyle's work, Stein and colleagues (Stein, Smith, Henningsen, & Silver, 2000; Stein, Grover & Henningsen, 1996) classify mathematical tasks according to **task features** and **level of cognitive demand**. Task features "refer to aspects of tasks that mathematics educators have identified as important considerations for the development of mathematical understanding, reasoning, and sense making" (Henningsen & Stein, 1997, p. 529), such as whether the task can be solved through a variety of strategies, through the use of multiple representations, and whether the task provides opportunities for mathematical communication, explanations, and justification. The level of cognitive demand refers to the type of thinking involved in solving the task. Resonant with Doyle's category of comprehension/understanding tasks, tasks that involve high levels of cognitive demand provide opportunities for students to engage in (1) "doing mathematics," complex thinking and reasoning such as exploring conjectures, forming generalizations, and justifying conclusions; or (2) "procedures with connections" to mathematical concepts, understanding, and meaning. Tasks that involve low levels of cognitive demands provide opportunities for students to engage in (1) "procedures without connections" to concepts or sense-making (i.e., Doyle's category of procedural or routine tasks); or (2) "memorization" of facts, formulae, or rules (i.e., Doyle's category of memory tasks). The Task Analysis Guide (TAG) presented in Figure 2.1 provides a complete

Lower-Level Demands	Higher-Level Demands
<u>Memorization Tasks</u> • Involve either producing previously learned facts, rules, formulae, or definitions OR committing facts, rules, formulae, or definitions to memory. • Cannot be solved using procedures because a procedure does not exist or because the time frame in which the task is being completed is too short to use a procedure. • Are not ambiguous – such tasks involve exact reproduction of previously seen material and what is to be reproduced is clearly and directly stated. • Have no connection to the concepts or meaning that underlay the facts, rules, formulae, or definitions being learned or reproduced. <u>Procedures Without Connections Tasks</u> • Are algorithmic. Use of the procedure is either specifically called for or its use is evident based on prior instruction, experience, or placement of the task. • Require limited cognitive demand for successful completion. There is little ambiguity about what needs to be done and how to do it. • Have no connection to the concepts or meaning that underlie the procedure being used. • Are focused on producing correct answers rather than developing mathematical understanding. • Require no explanations, or explanations that focus solely on describing the procedure that was used.	<u>Procedures With Connections Tasks</u> • Focus students' attention on the use of procedures for the purpose of developing deeper levels of understanding of mathematical concepts and ideas. • Suggest pathways to follow (explicitly or implicitly) that are broad general procedures that have close connections to underlying conceptual ideas as opposed to narrow algorithms that are opaque with respect to underlying concepts. • Usually are represented in multiple ways (e.g., visual diagrams, manipulatives, symbols, problem situations). Making connections among multiple representations helps to develop meaning. • Require some degree of cognitive effort. Although general procedures may be followed, they cannot be followed mindlessly. Students need to engage with the conceptual ideas that underlie the procedures in order to successfully complete the task and develop understanding. <u>Doing Mathematics Tasks</u> • Require 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). • Require students to explore and to understand the nature of mathematical concepts, processes, or relationships. • Demand self-monitoring or self-regulation of one's own cognitive processes. • Require students to access relevant knowledge and experiences and make appropriate use of them in working through the task. • Require students to analyze the task and actively examine task constraints that may limit possible solution strategies and solutions. • Require considerable cognitive effort and may involve some level of anxiety for the student due to the unpredictable nature of the solution process required.

Figure 2.1. The Task Analysis Guide (Stein et al., 2000).

description of Stein, et al.'s (2000) levels of cognitive demand of mathematical tasks. In the current study, a rubric based on the TAG will be used to assess the instructional tasks used by teachers in the study.

Analyzing Task Implementation. Doyle's work (1983, 1988) informs Stein and colleague's second premise for using instructional tasks as the basis for analyzing instructional episodes: tasks can exist at several different phases, and the cognitive demands of a task can potentially be altered during each of these phases. Doyle describes the phases as (1) the task as announced by the teacher, (2) the task as interpreted by the students, and (3) the task as reflected in the products expected by the teacher. These phases are all situated at the beginning of an instructional episode and address the potential cognitive processes that the task can elicit. Marx and Walsh (1988) condense Doyle's three phases of academic work into one phase entitled "task conditions," and extend Doyle's focus on potential cognitive processes to include phases that address the actual implementation of the task during instruction. Similarly, Stein and colleagues (1996) extend Doyle's phases, both forward and backward, to describe instructional tasks as passing through three phases: (1) tasks as they appear in print, before being announced by the teacher; (2) tasks as they are set up by the teacher (Doyle's first and third phases); and (3) tasks as they are enacted (i.e., carried out or worked on) by students and the teacher during the lesson. The last phase, referred to as the enactment or implementation of the task, extends Doyle's notion of the task as interpreted by students and the products expected by the teacher to encompass the cognitive processes actually performed and the products actually created by students through their work on the task. This focus on task implementation is consistent with the phases of academic work described by Marx and Walsh (1988). Stein and colleagues conclude their framework with a consideration of students' learning – the cognitive processes and features

of performance (i.e., what students know and can do) that result from engaging with the task (Stein & Lane, 1996).

To guide the analysis of classroom observations based on these phases, Stein and colleagues (1996, 2000) developed the Mathematical Tasks Framework (MTF) shown in Figure 2.2. The MTF explicates the relationship between instructional tasks and students' learning. Each phase represents segments of a lesson in which the cognitive demands of an instructional task are likely to be altered. In their analyses of classroom observations, QUASAR researchers identified the level of cognitive demand of the instructional tasks as set-up by the teacher and as implemented by the teacher and students during the lesson. In addition to the levels of cognitive demand identified in the TAG (Figure 2.1), two additional categories emerged during data analysis of task implementation: (1) non-mathematical activity and (2) unsystematic and/or nonproductive exploration. The category of unsystematic/nonproductive exploration was coded for the task implementation phase when students earnestly engaged with high-level cognitive processes but did not engage with the mathematical ideas embedded in the task (Stein & Lane, 1996).

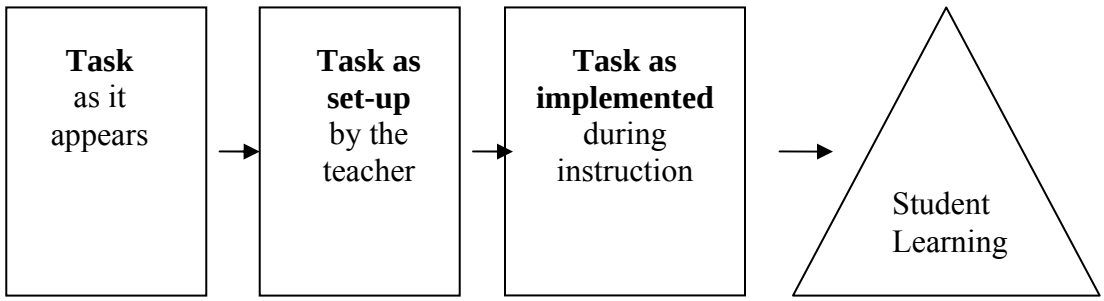


Figure 2.2. The Mathematical Tasks Framework (Stein et al., 2000).

The analyses identified classroom-based factors that influenced the maintenance and the decline of high-level cognitive demands as the task passed through the phases of the MTF and determined that specific sets of factors were associated with different patterns of enactment of instructional tasks (Henningsen & Stein, 1997). The classroom-based factors are provided in Figures 2.3. In this study, classroom observations will be analyzed using a rubric based on the level of cognitive demand of the task at each phase of the MTF, and the classroom-based factors will be used to provide qualitative descriptions of the features of instruction that supported or inhibited students' opportunities to engage with high-level cognitive processes during the observed instructional episodes.

The analyses of classroom observations in QUASAR yielded two major findings that have implications for the current investigation: (1) mathematical tasks with high-level cognitive demands were the most difficult to implement well, frequently transformed into less demanding tasks; and (2) student learning gains were greatest in classrooms in which high-level demands were consistently maintained and least in classrooms in which tasks were consistently of a procedural nature (Stein & Lane, 1996). These findings will be used in the next section to justify exposing teachers to the QUASAR frameworks as a basis for the professional development intervention in the current study.

2.3.2. QUASAR Frameworks as Professional Development and Research Tools

QUASAR researchers used instructional tasks and the nature of students' engagement with those tasks to understand the relationship between teaching and learning in project classrooms. This conceptualization situates mathematical tasks "in the interactions of teaching and learning" (Stein, et al., 2000, p. 25). Tasks provide the foundation for instruction, and other aspects of teaching and learning, such as opportunities for problem-solving and communication,

depend upon the features and cognitive demands of instructional tasks (Hiebert, et al., 1997; Doyle, 1988). As summarized by Doyle (1988), different kinds of tasks lead to different types of instruction, which subsequently lead to different opportunities for students' learning. Instruction that engages students with high-level cognitive processes (i.e., the type of cognitive processes that develop students' understanding of mathematics) is built upon challenging mathematical tasks (Hiebert, et al., 1997; Stein & Lane, 1996; Hiebert & Wearne, 1993).

Teachers influence tasks, and thus students' opportunities for learning, by defining and structuring the work that students do during instruction (i.e., determining the processes and resources that students are to use to accomplish the task and the products expected to result from students' work) (Doyle, 1988). Selecting worthwhile instructional tasks has been identified by researchers and teacher educators as an essential role of the teacher for promoting learning mathematics with understanding (Hiebert, et al., 1997; NCTM, 2000, 1991; Van de Walle, 2004). With the teacher as the agent who selects tasks and sets the parameters for how tasks will be enacted by students, teachers need to be aware of how different types of tasks influence students' opportunities for learning and how they can support students' engagement with high-levels cognitive processes during instruction. If the role of teachers is to facilitate conceptual understanding, then the first step in the process is for the teacher to select cognitively challenging mathematical tasks (Hiebert, et al., 1997). For these reasons, analyzing instruction based upon the cognitive demands of mathematical tasks and of the enactment of the tasks during instructional episodes is an essential and worthwhile focus for the professional development of teachers of mathematics.

Factors Associated with the Decline of High-level Cognitive Demands

1. Problematic aspects of the task become routinized (e.g., students press the teacher to reduce the complexity of the task by specifying explicit procedures or steps to perform; the teacher "takes over" the thinking and reasoning and tells students how to do the problem).

2. The teacher shifts the emphasis from meaning, concepts, or understanding to the correctness or completeness of the answer.

3. Not enough time is provided to wrestle with the demanding aspects of the task or too much time is allowed and students drift into off-task behavior.

4. Classroom management problems prevent sustained engagement in high-level cognitive activities.

5. Inappropriateness of tasks for a given group of students (e.g., students do not engage in high-level cognitive activities due to lack of interest, motivation or prior knowledge needed to perform; task expectations not clear enough to put students in the right cognitive space.

6. Students are not held accountable for high-level products or processes (e.g., although asked to explain their thinking, unclear or incorrect student explanations are accepted; students are given the impression that their work will not "count" toward a grade).

Factors Associated with the Maintenance of High-level Cognitive Demands

1. Scaffolding of students' thinking and reasoning.

2. Students are provided with means of monitoring their own progress.

3. Teacher or capable students model high-level performance.

4. Sustained press for justifications, explanations, and/or meaning through teacher questioning, comments, and/or feedback.

5. Tasks build on students' prior knowledge.

6. Teacher draws frequent conceptual connections.

7. Sufficient time to explore (not too little, not too much).

Figure 2.3. Factors associated with the maintenance and decline of high-level cognitive demands (Stein, Grover, & Henningsen, 1996).

Cognitive Demands of Instructional Tasks. Research has shown that teachers typically do not analyze tasks in terms of the type or level of thinking that the task can elicit from students. (Arbaugh & Brown, 2002; Stein, Baxter & Leinhardt, 1990). Studies conducted by Stein, Baxter, and Leinhardt (1990) and Arbaugh and Brown (2002) found that teachers categorized tasks with respect to similarities in mathematical content or surface-level features such as "word problems" or "uses a graph." Selecting instructional tasks based on levels of cognitive demand is not a common element of many teachers' mathematics pedagogy. Swafford and colleagues (Swafford, et al., 1997) found that teachers could design or alter lesson activities to contain features consistent with reform-oriented pedagogy (i.e., allow for student exploration or discussion), though teachers did not make changes that increased the cognitive demands of the instructional tasks. Even after teachers' participation in a course in which they were exposed to the van Hiele levels of geometric thinking, only 3% of the tasks in the overall post-lesson plans had the potential to elicit high-level cognitive demands (i.e., van Hiele Level 3 or higher). Other research consistently indicates that teachers' selection of instructional tasks is largely based on lists of skills and concepts they need to cover (Hiebert, et al., 1997) or by adhering to the tasks in their textbooks (Remillard, 1999; Doyle, 1983). Rather than thoughtful consideration of tasks that develop students' mathematical understanding, teachers often rely on instructional materials to provide and sequence instructional tasks. In classrooms observed by Doyle (1983), "academic work was defined in large measure by commercially available materials" (p.187) – curricular materials in which mathematics appeared as a set of discrete and interchangeable skills, to be practiced and mastered independently, with no logical or semantic thread connecting tasks or lessons from day to day, and an emphasis on computational accuracy and fluency rather than on

concepts or problem-solving. Doyle's findings are consistent with descriptions of U.S mathematics curricula in general (USDE-NCES, 2003; NCTM, 2000).

In this study, teachers will engage in analyzing the cognitive demands of mathematical tasks according to the Task Analysis Guide (TAG) (Stein, et al., 2000) presented in Figure 2.1. This taxonomy has been used in professional development settings (Smith, Stein, Arbaugh, Brown, & Mossgrove, 2004; Smith & Stein, 1998) and in teacher-development research (Arbaugh & Brown, 2002; Arbaugh, 2000; Smith, 1995). In a study conducted by Arbaugh and Brown (2002), the authors found that many teachers categorized mathematical tasks according to cognitive demands or opportunities for students' thinking after participating in a study group explicitly focused on the TAG and the Mathematical Tasks Framework (Figure 2.2). Interviews, journal entries, and collections of tasks provided evidence that the teachers attended to cognitive demands in the selection of instructional tasks in their own classrooms. Enhancing teachers' knowledge of the cognitive demands of mathematical tasks appears to be a valuable endeavor for professional development of teachers of mathematics because it provides teachers with a framework for assessing the tasks they select for instruction and the opportunities for students' learning of mathematics based on those tasks.

Implementation of Cognitively Challenging Tasks. Research has shown that implementing cognitively challenging tasks in ways that maintain students' opportunities to engage in high-level cognitive processes is not a trivial feat (USDE-NCES, 2003; Henningsen & Stein, 1997). Teachers often minimize task demands by breaking the task down into smaller sub-tasks (Smith, 2000), by focusing on correct answers and procedures (Henningsen & Stein, 1997; Romanagno, 1994; Doyle, 1988) or by adapting the tasks or teaching suggestions to be consistent

with their personal notions of effective teaching and learning (Remillard, 1999; Lloyd & Wilson, 1998; Clarke, 1997).

According to Smith (2000), an "important question for teacher educators is how to create experiences that help teachers build the capacity to support students as they struggle without reducing the cognitive demands of tasks" (p. 373). The present study will incorporate the Mathematical Tasks Framework (Figure 2.2; adapted from Stein, Grover, & Henningsen, 1996) as a tool for discussing, analyzing, and reflecting on the maintenance of high-level cognitive processes throughout an instructional episode. The "Factors Associated with the Maintenance and Decline of High-level Cognitive Demands" in Figure 2.3 will be used to engage teachers in analyzing and reflecting on instructional episodes and deriving general lessons learned from specific instructional cases of mathematics teaching. Teachers will also be exposed to research on the influence of the level of cognitive demand of instructional tasks on students' learning of mathematics (e.g., USDE-NCES, 2003; Stein & Lane, 1996).

The QUASAR frameworks were developed to analyze instructional episodes for research and have subsequently been used as tools to help teachers analyze instruction, reflect on their own practice, and discuss instruction with colleagues based on a shared language. There is no specific research to date indicating that exposing teachers to the MTF and the factors and patterns derived from QUASAR research will effectively support teachers' learning and instructional change—in fact, the intention of this study is to analyze the effectiveness of professional development based on the QUASAR frameworks in catalyzing teachers' learning and instructional change. Several professional development projects do, however, provide evidence that sharing research frameworks on students' mathematical thinking and how to support students' thinking during instruction can impact teachers' knowledge, beliefs and/or

instructional practices (Swafford, et al., 1997; Simon & Shifter, 1991; Carpenter, et al., 1989). Research also indicates that teachers implement tasks based on their own conceptions of effective mathematics teaching and learning (Smith, 2000; Remillard, 1999; Lloyd & Wilson, 1998; Clarke, 1997). Hence, enhancing teachers' knowledge of the cognitive demands of mathematical tasks and of the implementation of cognitively challenging tasks has potential to catalyze change in teachers' instructional practices with respect to the selection and implementation of high-level tasks in their own classrooms. The next section will review theories of learning to further support the argument that the intervention in this study will promote teacher learning in ways that will foster instructional change.

2.4. Theoretical Perspectives on Teachers' Learning

Prevalent in current frameworks for describing teachers' learning from professional development experiences are analogies between students' learning in ways that promote mathematical understanding and teachers' learning in ways that promote instructional change (Farmer, et al., 2003; Simon, Tzur, Heinz, Kinzel, & Smith, 2000; Cobb, Yackel, & Wood, 1991; Simon & Shifter, 1991). In this section, theories from cognitive psychology and social anthropology that describe students' learning of mathematics will be used to describe teachers' learning from professional development experiences (Cobb, Yackel, & Wood; 1991; Simon & Shifter, 1991; Borko & Putnam, 1995). Research on adults' learning will be reviewed to substantiate the analogy between teachers' learning and students' learning. The section will close by explicating the social-constructivist perspective on teachers' learning that served as the basis for structuring teachers' opportunities for learning in the present investigation.

2.4.1. Cognitive and Social Theories of Learning Mathematics

In mathematics, cognitive psychologists describe *learning* as the active construction of mathematical knowledge and *understanding* as the internal representation and structuring of mathematical ideas, connected in ways that allow the individual to recognize relationships between mathematical ideas and within different representations of a mathematical idea (Putnam, Lampert, & Peterson, 1990; Greeno, 1991; Hiebert & Carpenter, 1992). From cognitive psychology stems the constructivist perspective on teaching and learning mathematics prevalent in current efforts to reform mathematics education (NCTM, 2000). Based on the work of Piaget, the basic tenet of constructivism is that students construct mathematical knowledge as they interpret and attempt to solve the mathematical tasks encountered during mathematics instruction (Cobb, Yackel, & Wood, 1991; Simon & Shifter, 1991; Borko & Putnam, 1995). This process of constructing mathematical knowledge is stimulated when the individual encounters a problematic situation that creates disequilibrium with the individual's current knowledge structures. Hence, cognitively challenging tasks can serve to mediate an individual's opportunities to learn (Doyle, 1983). New mathematical knowledge and understandings are formed as the individual works to resolve the problem, resulting in a modification of the individual's previous knowledge structures and representations. Prior knowledge structures thus serve as the basis for the construction of new knowledge.

Sociological (or social anthropological) perspectives on students' learning of mathematics, as summarized by Voigt (1994), contrast cognitive psychology's focus on the individual's construction of knowledge by considering mathematical understanding as the product of social interactions. Based on the work of Vygotsky, *learning* is considered as the

internalization of social relationships, and *understanding* emerges from the shared, negotiated mathematical meanings created between students (rather than as the internal cognitive constructions of individuals) (Voigt, 1994). In shifting the focus to the social interactions of the community of learners and away from the individual student, the sociological perspective represents a view of learning antithetical to constructivism. Constructivism, referred to by Voigt (1994, p. 291) as "individualism," maintains that the student actively constructs mathematical knowledge, while the sociological view (referred to by Voigt as "collectivism") holds that mathematical knowledge exists as cultural and social practices in which the student is a participant.

Cobb, Yackel, and Wood (1991) contend that focusing only on the individual's internal construction of knowledge does not account for what can be seen happening as students work together in mathematics classrooms, and that focusing only on the community of learners does not explain students' individual insights or epiphanies that occur when working alone. Coordinating the two perspectives allows for both the individual student's private sense-making activities and for the ways of knowing developed though interactions with other students and through participation in the culture of learning mathematics in the classroom. This perspective on learning is referred to as social constructivism, and offers an explanation of how students' learning of mathematics is both interactive and constructive. It is often the interactions of groups of students that promote, maintain, and provide resolution to the conflicts that initiate individual students' construction of new mathematical knowledge (Cobb, Yackel, & Wood, 1991; Simon & Shifter, 1991).

Validating the Analogy between Students' Learning and Teachers' Learning. The theories from cognitive psychology and social anthropology presented in the previous section can be used, specific to the purposes in this investigation, to describe how teachers learn (Simon & Shifter, 1991). However, applying these perspectives to teachers' learning without consideration of the implications of teachers as adult learners would be remiss. Research on adult learning raises important similarities *and* differences between how children learn and how adults learn that lends further support to a social-constructivist approach to teacher professional development.

Rogers (2003) contends that while there is not an essential difference between the ways in which adults learn and the ways in which children learn, there are important differences in teaching adults and teaching children resulting from the identity of the adult-student versus that of the child-student. Other research in adult learning has identified self-concept or identity as a critical difference in adult and children's learning (Lai, 1995). The relationship between the teacher and the child-student in learning situations is far more consistent with their relationship in general than that of the teacher/facilitator and the adult-student. Adults must be given autonomy in their own educational experiences, opportunities to use their own judgments, and responsibilities for their own learning. Rogers does not characterize adults as a homogenous body of learners, but alternatively asserts that an individuals' experience in a learning situation will vary according to the person's prior experiences and their expectations for their own learning. Adults, more so than children, are inclined to learn only what they perceive to be meaningful, serving an interest or purpose in their current situation. Their life experience factors substantially into their expectations and openness to the learning situation (Lai, 1995). Hence,

Rogers conclusion that adults' learning, to an even greater degree than children's learning, must build on prior knowledge and experiences is consistent with the tenets of constructivism.

Rogers' ideas also appear to be consistent with research on teaching adults and children in specific content domains. Perfetti and Marron (1998) found that adults with low literacy levels who were learning to read experience the same specific types of difficulties encountered by children who have difficulty reading. These researchers contend that reading instruction for adults should focus on reading practice, with comprehension strategies and phonological awareness embedded within or resulting from actual reading experiences rather than taught or drilled directly. In this way, the adults' expectation for the learning experience (i.e., reading) is always at the forefront of instruction, with comprehension and decoding strategies offered or reinforced at the exact moment when they were both applicable and necessary to the task at hand.

Reading instruction that consists of opportunities to engage in reading would also appear to provide the adult with a sense of control over their learning process, perhaps sparking an interest in the materials they engaged in reading in the learning situation. Perfetti and Marron indicate that adults and children learning to read are different in their goals and motivation, with adults' goals and motivation stemming from practical or self-fulfilling purposes while children's arise from teacher- or parent-pleasing tendencies. Lai (1995) also notes that adults' orientation toward learning and motivation to learn are derived largely from their current interest or need in gaining knowledge on a particular topic. Consistent with Perfetti and Marron's belief that adults learn to read by reading, Lai suggests that adults learn best through modes of instruction that incorporate participation and dialogue as opposed to teacher-centered, lecture-based approaches that ignore the identity of the adult-student as an adult.

In summary, adults and children's learning differs along several dimensions: self-concept or identity, significance of prior experience, and motivation to learn (Rogers, 2003; Lai, 1995; Perfetti & Marron, 1998). These differences support a constructivist approach to adults' learning, and thus can be used as a basis to explain and to design opportunities for teachers' learning from professional development experiences. Teachers, in their identity as adults, need opportunities to construct their own knowledge, to wrestle with new ideas that they feel are relevant to their current situation in ways that build on, challenge, and enhance their prior knowledge, beliefs, and instructional practices.

2.4.2. Applying a Social Constructivist Perspective to Teachers' Learning

The remainder of this section will apply a social constructivist perspective to describe the process of teachers' learning from professional development experiences in ways that promote instructional change. This perspective – an interaction of the constructivist and sociological views on learning – is increasingly serving as the conceptual framework for research on teachers' learning (e.g., Farmer, et al., 2003; Simon, et al., 2000; Smith, 2000; Simon & Shifter, 1991; Cobb, Yackel, & Wood, 1991). Three components of this perspective appear to have significant implications for teachers' learning: (1) the importance of building professional development experiences on teachers' prior knowledge and beliefs; (2) the assertion that change occurs as new conceptions of mathematics teaching and learning conflict with the teachers' prior knowledge and beliefs; and (3) the role of social interaction in stimulating and maintaining this type of conflict. Each of these components will be described in the sections that follow.

<u>The Role of Prior Knowledge and Beliefs</u>. Mathematics teachers enter professional development experiences with their own prior knowledge and beliefs about mathematics, about

students as learners of mathematics, and about mathematics pedagogy (Borko & Putnam, 1995). Previous convictions serve as filters through which teachers interpret and come to understand new knowledge and ideas encountered within the professional development experience (Simon, et al., 2000; Lloyd, 1999; Remillard, 1999; Borko & Putnam, 1995; Fennema & Franke, 1992). Analogous to students' construction of new mathematical knowledge by building on prior knowledge structures (Romberg & Carpenter, 1986), professional development experiences must connect to teachers' current ways of thinking and use these prior conceptions as the basis for constructing new ideas about effective mathematics teaching and learning. If new ideas are either too disparate or entirely absent from teachers' current ways of thinking, the teacher possesses no cognitive structures through which the new idea can be attached or interpreted, and substantive change is not likely to result (Smith, 2000). Without connections to teachers' prior knowledge and beliefs, new knowledge is structured as an "accumulation of various kinds of skills and knowledge of practical routines, uninformed by general principles" (Farmer, et al., 2003, p. 341). The new knowledge may be additive but is neither generalized nor transformative.

Smith (1995) assets that, "Teachers, like students who are learning mathematics, need an opportunity to construct their knowledge in ways that build on what they bring to the experience and on their own ways of thinking and knowing (p. 23)." Pragmatically, this might be accomplished by providing opportunities in which teachers begin to see their current instructional practices as problematic (e.g., Smith, 2000; Cooney, 1999; Romanagno, 1994; Cobb, Yackel, & Wood, 1991), begin to gain new insights into students' mathematical thinking (e.g., Sherin, 2002; Wood, 1995; Carpenter, et al., 1989), or enhance their own understanding of mathematics and reflect on their personal learning experiences (e.g., Farmer, et al., 2003; Simon

& Shifter, 1991). Simon and colleagues (Simon, et al., 2000) provide the following explanation for the influence that a teachers' current conceptions have on their learning:

> "First, …what a learner perceives (attends to) and the interpretations that she makes are structured by her current knowing (Piaget's notion of assimilation). Second, it is this current knowing that is transformed (Piaget's notion of accommodation). Thus, the possibilities for learning are afforded and constrained by the current state of knowing (p. 584)."

Similarly, Borko and Putnam (1995) contend that the knowledge and beliefs that filter and interpret new ideas are the same structures that are the targets of change. How then, does change occur?

Conflict as Impetus for Instructional Change. Most teachers will act in ways consistent with their beliefs about how mathematics is best taught and learned (Cooney, 1999; Remillard, 1999; Borko & Putnam, 1995; Thompson, 1992). These beliefs are questioned as new ideas about supporting students' mathematical learning arise in professional development settings (Borasi, et al., 1999; Simon & Shifter, 1991; Carpenter, et al., 1989), in attempts to implement new curricular materials (Remillard, 1999; Lloyd, 1999; Lloyd & Wilson, 1998; Clarke, 1995) or as instructional dilemmas arise in implementing reform-oriented instruction in their own classrooms (Smith, 2000; Romanagno, 1994; Cobb, Yackel, & Wood, 1991). The construction of new understandings is stimulated by a problem situation or disequilibrium that occurs when teachers' current cognitive structures cannot simply be refined or supplemented to assimilate the new idea. According to Simon & Shifter (1991), disequilibrium leads to changes in previously held ideas and convictions to account for the new experience. Smith (1995) argues that when a new conception of teaching is in sharp contrast with current beliefs, "the teacher begins the

process of trying to eliminate the disequilibrium" (p. 38) that the conflict has created. Resolving the conflict and reestablishing equilibrium promotes change by initiating a reorganization of the teacher's existing knowledge and beliefs in order to accommodate the new conception.

Once the new information has created a conflict which the teacher seeks to resolve, changes in current knowledge and beliefs initiate changes in instructional practices to reflect the new insights into mathematics pedagogy or students' thinking. It would seem rather antithetical for teachers to maintain instructional practices that they had come to believe did not support their students' learning (Cooney, 1999). In providing teachers with experiences that cause them to question previously held conceptions, teachers will begin to modify current beliefs and instructional practices to accommodate new ideas. Hence, in order to effect instructional change, professional development experiences must initiate doubt that a teacher's current modes of instruction are best suited to foster students' learning, or create conflict between insights gained from the new experiences and previous conceptions of mathematics, students as learners of mathematics, or effective mathematics pedagogy. This induction of "doubt" is an essential catalyst for stimulating inquiry and reflection, rejecting prior knowledge and beliefs, and ultimately changing one's teaching practice in accordance with new knowledge and beliefs (Cobb, Yackel, & Wood, 1991; Wallen & Williams, 2000). Conflict or doubt initiates transformative learning (Thompson & Zeuli, 1999) or a stance of inquiry (Farmer, et al., 2003) that is capable of catalyzing instructional change.

Role of social interactions. In the social constructivist view of teacher learning, changes in teachers' beliefs and practices are thought to result from an interaction of the cognitive conflict and social-interactionist perspectives on learning. A cognitive conflict perspective asserts that opportunities for learning occur as teachers reorganize their experiences to resolve

conflicting viewpoints and consider new courses of action (Cobb, Yackel, & Wood, 1991; Simon & Shifter, 1991). Because other members of the group are instrumental in initiating and maintaining these conflicts, social constructivist theories support the contention that the interaction of participants during the professional development experience is a crucial component in challenging teachers' existing knowledge and beliefs and promoting change. A new experience may initiate conflict within an individual by raising issues that contrast with a teacher's current beliefs and practices; however, these issues can remain unchanged and unresolved in the absence of colleagues to challenge the position or offer alternative actions or ways of thinking (Wallen & Williams, 2000; Simon & Shifter, 1991). For example, a teacher may consider a pedagogical move to be problematic (or conversely, to be particularly effective) without deep consideration of why or of what effect the move might have had on students' opportunities to learn mathematics. Though an issue may resonate with an individual's personal struggle, simply acknowledging a connection with the issue without the input and guidance of others is not likely to provide a pathway toward resolution.

Interaction provides a forum for maintaining conflicts in productive ways, first, as the individual is pressed to discuss critical issues and defend their position to others and second, for initiating new conflicts as the individual engages in challenging the contentions of other participants. The interaction of the group allows the conflict to become an impetus for cognitive growth and change in individual participants. Hence, Cobb, Yackel, and Wood (1991) describe learning as both an interactive and constructive activity. Accommodations to the knowledge structures of individuals are influenced by the ideas of other participants as the group negotiates their understandings of best practices in mathematics teaching and learning. Concurrent and interdependent with the individual's active engagement in constructing their own understandings

is the social construction of knowledge by the group of learners to which the individual is a member (Voigt, 1994; Simon & Shifter, 1991). Opportunities for teacher learning are optimized when teachers become members of a community of learners in which social norms govern the rules of interacting and the expectations for the level and type of engagement. A critical component for the norms of the group discussion and interaction is to cast conflicts, disagreements or issues with which individuals are struggling as opportunities for learning. In this way, group discussions become situations where teachers are pressed to deeply consider their initial position with respect to the viewpoints or alternative approaches suggested by others (Barnett, 1998; Simon & Shifter, 1991).

The following section will situate the intervention in the current study within the research and theories of effective professional development presented throughout this chapter.

2.5. Framing the Current Study

This section draws on the research and theories of transformative professional development and teacher learning presented throughout the chapter in order to situate the current investigation within the current best thinking on professional development for teachers of mathematics. As a closing to Chapter 2, this final section argues that the intervention in the present study will be effective in promoting teacher learning and instructional change and that the analysis proposed in Chapter 3 will be a reasonable and effective means of assessing the learning and instructional change of teachers participating in the investigation.

2.5.1. Why will the Intervention be Effective?

The intervention in the current investigation consists of a set of professional development experiences for mathematics teachers that intends to foster changes in teachers' instructional practices; specifically, in teachers' selection and implementation of cognitively challenging tasks for mathematics instruction in their own classrooms. The selection and implementation of cognitively challenging tasks has been established as an important and worthwhile focus for teachers' learning because teachers select the tasks that students engage with during mathematics instruction, ands tasks structure students' opportunities to learn mathematics (Stein & Lane, 1996; Stein et al., 1996; Hiebert & Wearne, 1993; Doyle, 1983). Basing professional development on research frameworks developed by the QUASAR Project will provide opportunities for teachers to enhance their knowledge and instructional practices with respect to the selection and implementation of cognitively challenging mathematical tasks (Arbaugh & Brown, 2002; Stein et al., 2000).

Analogous to the influence of mathematical tasks on students' learning, the tasks that teachers engage with during professional development experiences structure their opportunities to learn mathematics pedagogy. Earlier sections in this chapter validated the analogy between students' and teachers' construction of knowledge. Students actively construct mathematical knowledge as they interpret and attempt to solve cognitively challenging mathematical tasks. Cobb, Yackel, & Wood (1991) assert that "the constructivist dictum that knowledge is constructed by reorganizing experiences to resolve problems applies as much to teachers as it does to students" (p. 89). Just as direct teaching frequently leads to rote learning of mathematics, the content, ideas, and principles that teachers encounter in professional development settings are unlikely to have an effect on their current knowledge and instructional practices if this

information is delivered as the transmittal of formal knowledge and theories of effective teaching and learning (Ball & Cohen; 1999; Cobb, Yackel, & Wood, 1991; Simon & Shifter, 1991). Change will not happen by teachers passively absorbing new information through lecture (Ball & Cohen, 1999; Thompson & Zeuli, 1999). The intervention in this study is grounded in social-constructivist theories of learning, and, as recommended throughout the professional development literature, models the types of instructional practices that are intended for teachers to implement with their students (Farmer, et al., 2003; Borasi, et al., 1999; Ball & Cohen, 1999; Simon & Schifter, 1991). Teachers participating in the current investigation will be provided with opportunities to construct new knowledge by interpreting and attempting to resolve issues that challenge their current conceptions of effective mathematics pedagogy. Both small and large group discussions will be utilized to initiate and maintain conflicts and to create a community where teachers can resolve these conflicts and negotiate new understandings of mathematics, mathematics pedagogy, and students' learning (Barnett, 1998; Cobb, Yackel, & Wood, 1991). Through group discussions, teachers will refine their knowledge and beliefs by publicly articulating their own thoughts, making sense of the ideas and perspectives of their colleagues, and by resolving the disequilibrium created when individuals hold different viewpoints or perspectives (Simon & Shifter, 1991; Borasi, et al., 1999). The professional development experiences will be closely aligned with issues relevant to the teachers' current situation in order to encourage reflection that is oriented towards the teacher's own practice (Wallen & Williams, 2000; Ball & Cohen 1999).

In line with recommendations by Ball and Cohen (1999), an underlying theoretical framework guided the design and selection of the activities used within and across the professional development sessions that form the treatment in this investigation. This framework

focuses on the cognitive demands of mathematical tasks, the phases during a lesson at which the cognitive demands of mathematical tasks are likely to be altered, and the classroom factors that serve to influence the maintenance or decline of high-level cognitive demands (Stein, et al., 1996; Henningsen & Stein, 1997). Specifically, the Mathematical Tasks Framework (MTF) (see Figure 2.2) provided a coherent thread that situated the activities and goals for teacher learning within individual sessions and across the set of sessions. For example, in Session 1, teachers engaged in solving two mathematical tasks, analyzing the different opportunities for students' learning provided by the tasks, and sorting a set of mathematical tasks based on cognitive demands. Session 1 focused almost exclusively on the levels of cognitive demand and the influence of mathematical tasks on students' learning. Following Session 1, the project teachers were asked to read about two different instructional episodes (i.e., a dual case) featuring one of the tasks that they had solved during the session. Session 2 began with a discussion comparing students' opportunities for learning in each of the instructional episodes, with participants identifying classroom factors that influenced the maintenance of high-level cognitive demands in one lesson and the decline of high-level cognitive demands in the other.

The professional development sessions also provided opportunities for project teachers to examine, analyze, and reflect on mathematics, mathematics pedagogy, and student thinking by exploring artifacts representative of the everyday work of teaching (Smith, 2001; Ball & Cohen, 1999; NCTM, 1991). According to Smith (2001), "In this view, materials that depict the work of teaching (e.g., student work mathematical instructional tasks, and classroom episodes) are used to create opportunities for critique, inquiry, and investigation" (p.2). Project teachers engaged in solving and analyzing mathematical tasks, analyzing and reflecting on instructional episodes, and assessing students' mathematical understanding evident during the instructional episodes and in

samples of student work. The goal in using these activities was to provide opportunities for project teachers to extract general theories or "lessons learned" that would catalyze changes in their own instructional practices – changes that would lead to the selection and implementation of high-level tasks in their own classrooms. Hence, activities used within the professional development sessions were selected and designed to represent "samples of authentic practice" (Smith, 2001, p.7) through which we could focus teachers' learning on the cognitive demands of mathematical tasks and the maintenance or decline of cognitive demands at each phase of the MTF.

Toward this purpose, the book, Implementing Standards-Based Instruction: A Casebook for Professional Development, (Stein, Smith, Henningsen, & Silver, 2000) was provided to project teachers. The book served both as a source of practice-based materials (i.e., mathematical tasks and cases of mathematics instruction) and as a resource for communicating critical elements of the underlying theoretical frameworks (i.e., levels of cognitive demand, the MTF, and the factors influencing the maintenance or decline of high-level cognitive demands) upon which the professional development sessions were based. These frameworks, and the ways in which they supported teachers' work throughout the sessions, were continually made explicit to teachers. Furthermore, teachers were provided with opportunities to assess their own instructional practices using the lens of the MTF and the factors that support and inhibit the maintenance of high-level cognitive demands. These opportunities included: identifying the cognitive demands of tasks in their own curriculum; reflecting on a lesson in which they based instruction on a high-level task; identifying factors for maintaining the cognitive demands of mathematical tasks that they intend to work on in their own classrooms, analyzing instructional episodes to assess their progress on these factors; analyzing student work for evidence of high-

level engagement; creating questions to support students' engagement with a high-level task; and planning a lesson with specific factors in mind.

Consistent with recommendations by researchers and teacher educators (e.g., NCTM, 2000, 1991; Ball & Cohen, 1999; Borko & Putnam, 1995; Simon & Schifter, 1991), the facilitators modeled the type of instructional strategies that were intended for project teachers to begin to incorporate into their own classrooms. Based on a social-constructivist view of teacher learning (Borko & Putnam, 1995; Simon & Schifter, 1991), these instructional strategies provided a supportive and collaborative environment in which project teachers were challenged to wrestle with new ideas, to openly express disagreements, and to identify aspects of their own instructional practices that they would like to change. In essence, the facilitators endeavored to foster the type of disequilibrium that generates changes in teachers' knowledge that subsequently catalyze changes in teachers' instructional practices. Another salient aspect of the professional development seminars were opportunities for exploration and for sharing ideas with colleagues and small- and whole-group discussions that pressed teachers to make important mathematical and pedagogical connections. These discussions encouraged and maintained teachers' engagement with issues and ideas that were likely to generate the types of conflicts and doubt of current knowledge and beliefs that lead to instructional change (Cobb, Yackel, & Wood, 1991). Aligned with views on the role of tasks in mediating teaching and learning, the professional development sessions engaged teachers with cognitively challenging professional learning tasks (i.e., tasks that involved analysis, reflection, and generalization [Smith & Stein, 2002]) and supported teachers' work in ways that provided opportunities for thinking, sense-making, and engagement with high-level cognitive processes.

2.5.2. Why will the Analysis be a Valid Means of Measuring Teacher Learning and Instructional Change?

This chapter has provided a review of professional development studies that inform the present investigation. Similar to the study being proposed here, these studies engaged teachers in interventions designed to catalyze changes in teachers' knowledge, beliefs, and practices, and subsequently analyzed the changes that occurred. Descriptions of the evidence of teacher learning and instructional change provided earlier in this chapter revealed that many of the studies relied on teachers' self-reports (teachers' writings, interviews, and surveys) to proclaim changes in teachers' knowledge, beliefs, and instructional practices (e.g., Farmer, et al., 2003; Borasi, et al., 1999). Furthermore, most of the studies that did collect classroom artifacts or conduct classroom observations did not analyze these sources of data in ways that were consistent with the goals of the study or of the professional development in the study (e.g., Swafford, et al., 1997) or did not analyze the observational data at all (e.g., Simon & Schifter, 1991). Evidence of changes in teachers' knowledge and beliefs is often justifiable; the studies provided teachers with professional development experiences that enhanced teachers notions of effective teaching and learning in mathematics, and these newly developed conceptions were apparent in teachers' self-reports in ways that could be reasonably attributed to the professional development interventions. However, many of the studies (1) make broad claims of changing teachers' instructional practices that were not based on classroom artifacts or observations; and (2) did not assess implementation.

In the present investigation, changes in teachers' knowledge of the cognitive demands of mathematical tasks and of ways of maintaining high-level cognitive demands during implementation will be considered as evidence of teacher-learning. Changes in the selection and

implementation of cognitively challenging instructional tasks in teachers' classrooms will be considered as evidence that their learning was transformative; in other words, as evidence of instructional change. This study will measure changes in teachers' knowledge and instructional practices using pre/post measures. The study will also make extensive use of classroom artifacts (i.e., instructional tasks and student work) and observations to measure changes in teachers' selection and implementation of cognitively challenging tasks. The present investigation builds on QUASAR frameworks as professional development tools and as tools to analyze teacher learning and instructional change in ways that reflect the goals of the study -- to influence teachers' selection and implementation of cognitively challenging mathematical tasks during instruction in their own classrooms. In this way, the analyses of changes in teachers' knowledge and instructional practices are designed to assess the specific focus and intended outcome of the professional development experiences. Collection of pre/post measures, utilization of observational data, and consistency between goals of the professional development and analysis of the data constitute ways in which the present investigation refines the research design of prior professional development studies reviewed earlier in this chapter.

Hence, the present investigation builds on current knowledge of transformative professional development and extends earlier efforts at analyzing instructional change. Next, Chapter 3 will provide a detailed description of the intervention and of the research design for the study.

3. CHAPTER 3: METHODOLOGY

This investigation determined the extent to which professional development experiences based on the selection and implementation of cognitively challenging mathematical tasks influenced the ways in which mathematics teachers selected and implemented mathematical tasks in their own classrooms. Changes in teachers' knowledge were assessed using pre- and post- measures and interviews focused on the cognitive demands of mathematical tasks and the implementation of tasks with high-level cognitive demands. Changes in teachers' instructional practices were assessed at three points during the school year in which they engaged in the professional development seminars (2004-2005) by rating the cognitive demands of instructional tasks, student work, and observed instructional episodes in the teachers' classrooms. Artifacts and other records of teachers' participation in the professional development seminars were used to form reasonable connections, though not causal links, between changes in teachers' knowledge and instructional practices and their experiences in the professional development sessions. Furthermore, data from a contrast group was used to gage whether the knowledge and instructional practices of project teachers differed from the knowledge and instructional practices teachers who did not participate in the professional development intervention.

Quantitative data (i.e., frequency and/or rubric scores) on the levels of cognitive demand of instructional tasks, student work, and classroom observations were analyzed for evidence of change over time using descriptive statistics. Qualitative research methods were used to describe

the nature of changes in teachers' selection and implementation of mathematical tasks, based on instructional factors and patterns known to influence the maintenance or decline of high-level cognitive demands (Stein, Grover, & Henningsen, 1996; Henningsen & Stein, 1997) and on themes and patterns emergent in the data (Janesick, 2000). This process reduced the data "into a compelling, authentic, and meaningful statement" of the type of changes in teachers' knowledge and instructional practices that occurred in the study (Janesick, 2000, p. 388). As recommended by experts in qualitative research, this investigation provided narratives of the experiences of individual teachers that characterized sets of teachers with similar patterns of change (Janesick, 2000). Hence, quantitative research methods and a pre/post design were used to determine whether changes occurred in teachers' knowledge and instructional practices, and qualitative methods were used to describe the nature of these changes and how they related to the professional development intervention.

Both triangulation and clarity of focus provided validity to the research design in this study (Denzin & Lincoln, 2000). The research design allowed for the triangulation of data and of research methods by providing multiple sources of evidence of changes in teachers' knowledge and instructional practices (Denzin & Lincoln, 2000; Janesick, 2000). Furthermore, the study maintained a clear focus on the cognitive demands of mathematical tasks throughout the research questions, the professional development intervention, and the collection and analysis of data. In this way, the research design provides a credible explanations and inferences to appropriately answer the research questions (Janesick, 2000). The remainder of this chapter describes the context of the study, the selection of subjects in the study, the data sources, and the procedure for analyzing the data.

3.1. Context of the Study

This investigation focuses on secondary mathematics teachers participating in a professional development project at the University of Pittsburgh, entitled Enhancing Secondary Mathematics Teacher Preparation (ESP). In order to provide a context for the study, the goals of the project and the design of the professional development sessions in which ESP teachers participated are described in the following section.

3.1.1. Goals of the ESP Project

The ESP Project was initiated in the fall of 2003 with the overarching goal of improving the preparation and field experiences of mathematics teaching candidates at the University of Pittsburgh. The ESP Principal Investigators (Dr. Margaret Smith and Dr. Ellen Ansell, School of Education, and Dr. Beverly Michaels and Dr. Paul Gartside, Mathematics Department) devised three components to accomplish this goal:

Component 1: The creation of two additional mathematics courses specifically targeted at making connections between formal mathematics courses and the mathematics that is at the heart of the secondary mathematics curriculum. These courses are intended to deepen teachers' understanding of the mathematics needed for teaching.

Component 2: The revision of the existing secondary mathematics methods courses to incorporate practice-based learning experiences (Smith, 2001) and current research on effective mathematics teaching and learning.

Component 3: The development of a cadre of mentor teachers who can enact, support, and promote mathematics education reform efforts in the school environments in which they work and who can provide support to pre-service teachers during their student internship

experiences and to new teachers within the mentors' schools.

Hence, the ESP project intended to impact mathematics education reform in the Pittsburgh region by enabling mentor teachers to serve as lead teachers and promote mathematics reform efforts within their schools, by enabling preservice teachers to become future leaders in mathematics education reform, and by improving the pool of teaching candidates in the region. As summarized by Smith (2003), "By improving the preparation of mathematics teachers, the ESP project seeks to improve mathematics *teaching* and thereby improve the mathematics *learning* of students (p. 39)."

3.1.2. The ESP Professional Development Workshop as the Intervention in this Study

To achieve the goals of Component 3, the ESP project team designed a set of professional learning experiences for mentor teachers. These experiences began with a professional development workshop focused on teaching and learning in the teachers' own classrooms. Mentor teachers attended this workshop during their first year of participation in the project. This workshop consisted of six one-day sessions, held on Saturdays throughout the school year (i.e., in October, November, January, February, March, and May). The first ESP workshop was conducted during the 2003-2004 and served as a pilot for the design and content of the 2004-2005 workshop that constitutes the "treatment" of the present investigation. Many of the professional learning tasks remained consistent for the 2004-2005 professional development workshop, though revised based on feedback and reflection from Cohort 1 teachers and the ESP development team. An overview of the professional learning tasks that collectively formed the content of the 2004-2005 professional development workshops is provided in Figure 3.1, and a summary of these activities is provided in Appendix 3.1.

Session 1: Oct. 2, 2004	Session 2: Nov. 6, 2004	Session 3: Jan. 8, 2005	Session 4: Feb. 5, 2005	Session 5: Mar. 5, 2005	Session 6: May 7, 2005
Introductions & Data Collection	Introducing Levels of Cognitive Demand and The Mathematical Tasks Framework	Reflecting on Sessions 1 & 2	Why Cases?	Case Stories II: Storytelling through Student Work. What did students' work tell about maintaining high-level cognitive demands during the lesson?	Case Stories III: How did assessing & advancing questions influence the enactment of the task?
Solving "Martha's Carpeting" & the "Fencing" Tasks	Solving the "Linking Fractions, Decimals, & Percents" Task	Multiplying Monomials and Binomials: Developing the area model of multiplication	Case Stories I: Reflecting on Our Own Practice. How did the factors of scaffolding and press play out in the lesson?		Planning the "Sharing and Discussing" Phase of a Lesson: Selecting and ordering presentations
Comparing Martha's Carpeting Task & the Fencing Task: How are they same and/or different?	Reading & Discussing the Case of Ron Castleman: Similarities and differences between 2nd and 6th period. Do the differences matter?	Solving the "Multiplying Monomials & Binomials" Task with Algebra Tiles	Solving the "Extend Pattern of Tiles" Task	Focusing on the "Exploring the Task" Phase of a Lesson: What questions would you ask to assess and to advance students' understanding?	Introducing the "Thinking Through a Lesson" Protocol
Categorizing Mathematical Tasks: The Task Sort	The Factors and Patterns of Maintenance & Decline	Reading & Discussing the Case of Monique Butler: What did MB *want* her students to learn and what *did* they learn?	Analyzing Student Work on the Extend Pattern of Tiles Task: Which show greatest/least understanding?	Planning a Whole-Group Discussion: What responses would you share & why?	Solving "Double the Carpet" Task
Data Collection, Paperwork	Data Collection, Paperwork	Connecting to Own Teaching: Discuss factors that influenced your lesson			Data Collection, Paperwork
Identify a task from your data collection that you would like to change/adapt/improve in some way.	*Plan, Teach and Reflect on a lesson involving a high-level task: identify factors at play in your lesson and factors you want to work on this year*	*Plan, teach and reflect a lesson using a high-level task. In what ways did you make progress on the factor you have chosen? What do you still need to work on?*	*Plan, Teach and Reflect on a lesson involving a high-level task: before and after, complete the chart on factors and expectations. Bring in student work.*	*Plan, Teach and Reflect on a lesson involving a high-level task. List questions to assess & advance Ss learning. Bring in list of questions and student work.*	*Plan, Teach and Reflect on a lesson involving a high-level task. Use the TTAL to plan and reflect on the "Sharing & Discussing" phase of your lesson.*

Figure 3.1. ESP professional development activities for Cohort 2 (2004-2005).

Following the ESP professional development workshop, mentor teachers participated in two additional professional learning experiences that are not under study in this investigation because they did not focus on the selection and implementation of instructional tasks in teachers' classrooms. During June 2005, mentor teachers participated in a one-week summer workshop focused on leadership and mentoring, designed to prepare mentor teachers to support the pre-service teacher assigned to their classroom. Over the 2005-2006 school year, mentor teachers attended 5 half-day sessions, accompanied by their pre-service teachers, in which they collectively engaged in analyzing and reflecting upon effective mathematics teaching and learning.

3.2. Selection of Subjects

This section describes the recruitment and selection of teachers to participate as subjects in the ESP group or as subjects in the contrast group.

3.2.1. Selection of the ESP Group

The secondary mathematics teachers participating as mentor teachers in Component 3 of the ESP Project during the 2004-2005 school year were asked to participate as the subjects in this study. These teachers formed the second cohort of ESP mentor teachers. The first cohort began the ESP Project in the fall of 2003, attended professional development sessions throughout the 2003-2004 school year and the summer of 2004, and mentored preservice teachers during the 2004-2005 school.

Teachers for the second cohort were recruited during the spring of 2004 by 1) directly contacting teachers who were interested but unable to participate in the first cohort; 2) having

teachers in the first cohort recruit colleagues from their schools; 3) contacting school administrators from districts who were interested in participating in the first cohort; and 4) contacting administrators from other schools in the region in which the University of Pittsburgh has traditionally placed pre-service teachers. In instances where school administrators were contacted, it was often the case that specific teachers were recommended and that the administrators were asked for their recommendations. Nineteen teachers were recruited to participate in the second ESP cohort and were given the option of participating in this study. The teachers were provided with a stipend of $1000 for participating in the ESP professional development workshop (during the 2004-2005 school year) and the leadership and mentoring workshop (during June 2005); receipt of the stipend was not contingent upon teachers' participation in this study. All nineteen teachers agreed to take the pre- and post-test and to provided data from the professional development sessions, eighteen teachers agreed to provide classroom artifacts (tasks and student work), and a subset of 11 of the 18 teachers also elected to participate in classroom observations.

Teachers participating in the ESP project have been chosen as subjects for this study for several reasons. First, all the teachers are from the local region and teach in schools that have relationships with the University of Pittsburgh, which provides the potential for accessibility to the teachers and their classrooms. The second cohort of mentor teachers has been chosen because the ESP professional development team could draw on experiences and feedback from working with the first cohort to make improvements to the structure and content of the professional development sessions. Finally, the second cohort was anticipated to be a larger group than the first, providing the potential for more subjects for the study.

Certain assumptions about the characteristics of subjects in this study can be made based on procedures for recruiting participants for the ESP project. First, ESP teachers must have at least three years teaching experience, with at least two years in their current school district, in order to mentor a student teacher according to state law. Second, the teachers recommended or specifically targeted for participation in the second ESP cohort were recognized by colleagues, school administrators, or ESP project team members as potential mathematics teacher leaders. Specific recruitment at a large urban school that incorporated reform-oriented curricula and engaged mathematics teachers in quality professional development opportunities provided a pool of potential subjects who were already in the process of incorporating reform-oriented teaching practices, such as selecting and implementing high-level tasks

The eighteen teachers participating in this study ranged from 3 to 30 years of teaching experience (12 teachers had less than 10 years experience, 4 teachers had 10-15 years experience, and 2 had over twenty years experience), with an average of 8.5 years in the classroom. At the time of their participation in ESP, 9 teachers in cohort 2 were teaching middle school mathematics and 9 were teaching high school mathematics. Seventeen of the teachers were certified to teach secondary mathematics, and the remaining teacher held a certification as an elementary teacher. School demographics spanned the range from a large, urban, economically disadvantaged school district to a mid-sized affluent suburban school to several small middle-class suburban schools. The teachers' professional development opportunities (outside of the ESP project) varied greatly, as did exposure to and use of reform-oriented mathematics curricula and ways of teaching mathematics. Some of the teachers in this study participated in other research on teachers' learning and instructional change conducted at the

University of Pittsburgh. In these instances, University of Pittsburgh researchers in the School of Education have maintained the same pseudonyms across studies.

3.2.2. Selection of the Contrast Group

A contrast group was selected to provide a means of assessing whether the knowledge and instructional practices of ESP teachers differed from the knowledge and instructional practices of teachers who did not participate in the ESP professional development workshop. Six school districts in the Pittsburgh region (but not participating in ESP) were contacted, and district administrators (i.e., the superintendent, curriculum director, and building principals) were provided with an overview of the study, the goals ESP professional development workshop, and the data to be collected from contrast group teachers. Two school districts agreed to recruit teachers to participate in the contrast group, and all mathematics teachers in the two schools were given the option of participating. The purpose of their participation in study was described to teachers as "helping to determine, at the end of the school year, if teachers who had participated in a year-long professional development workshop with a very specific focus had different knowledge or instructional practices (related to the specific focus) than teachers who had not participated in the professional development workshop." No stipend was offered to contrast group teachers, though administrators and teachers in both schools requested (and received) consulting and/or professional development activities in exchange for their participation.

A total of 10 teachers (five from each school) agreed to serve as contrast subjects. The ten teachers ranged from 2 to 31 years of experience (4 teachers with less than 10 years experience, 3 with 10-15 years experience, and 3 with over 20 years experience), with an average of 11.8 years in the classroom. At the time of the study, 6 teachers in the contrast group

were teaching middle school mathematics and 4 were teaching high school mathematics. All 10 teachers held certification to teach secondary mathematics. Both school districts were located in suburban areas, with one school serving an affluent community and the other serving a middle-class community. One school was implementing reform-oriented curricula in middle school and high school and was participating in a large-scale professional development project for teachers of mathematics. The other school did not use reform-oriented curricula and did not offer quality professional development opportunities for mathematics teachers. Hence, the schools and teachers in the contrast group reflect the diversity of the ESP group along several dimensions: years of teaching experience, teaching middle vs. high school, school demographics, variation in professional development opportunities for teachers, and variation in use of reform-oriented mathematics curricula.

3.3. Data Sources

Data collection began during the first ESP professional development session in October 2004, with a pre-test that was administered to participants. At three points during the 2004-2005 school year, a data-set was collected from each ESP teacher that consists of: (1) the instructional tasks used in the teacher's classroom over a five day period, (2) student work from a subset of three of these tasks, and (3) one classroom observation within the same 5-day period (see Figure 3.2). These data-sets will be referred to as the Fall (October 2004), Winter (January 2005), and Spring (May 2005) data collections. Data from teachers in the contrast group was collected during the spring of 2005 and consists of (1) the same pre-test used with ESP teachers, and (2) one classroom observation. Data on ESP teachers' participation and experiences in the professional development sessions consists of videotaped records and collections of artifacts

from the sessions. Finally, a post-test and post-ESP interview was conducted at the conclusion of the professional development seminars (June, 2005). Figure 3.3 provides a timeline of the ESP data collection. Each of the data sources will be described in detail in the following section.

3.3.1. Pre- and Post-Test Task Sort

The instrument used as the pre- and post-test is a written-response card sort, with each card containing one mathematical task (hereafter referred to as the ***task sort***). The purpose of the task sort was to provide an assessment of teachers' pre- and post-knowledge of the cognitive demands of mathematical tasks. The design, use, and analysis of the task sort in the present investigation was informed by research establishing the use of card sorts to elicit teachers' knowledge (Stein, Baxter & Leinhardt, 1990) and by research specifically focused on using a task sort to assess teachers' ability to differentiate between tasks with high-level and low-level cognitive demands (Arbaugh & Brown, 2002; Arbaugh, 2000).

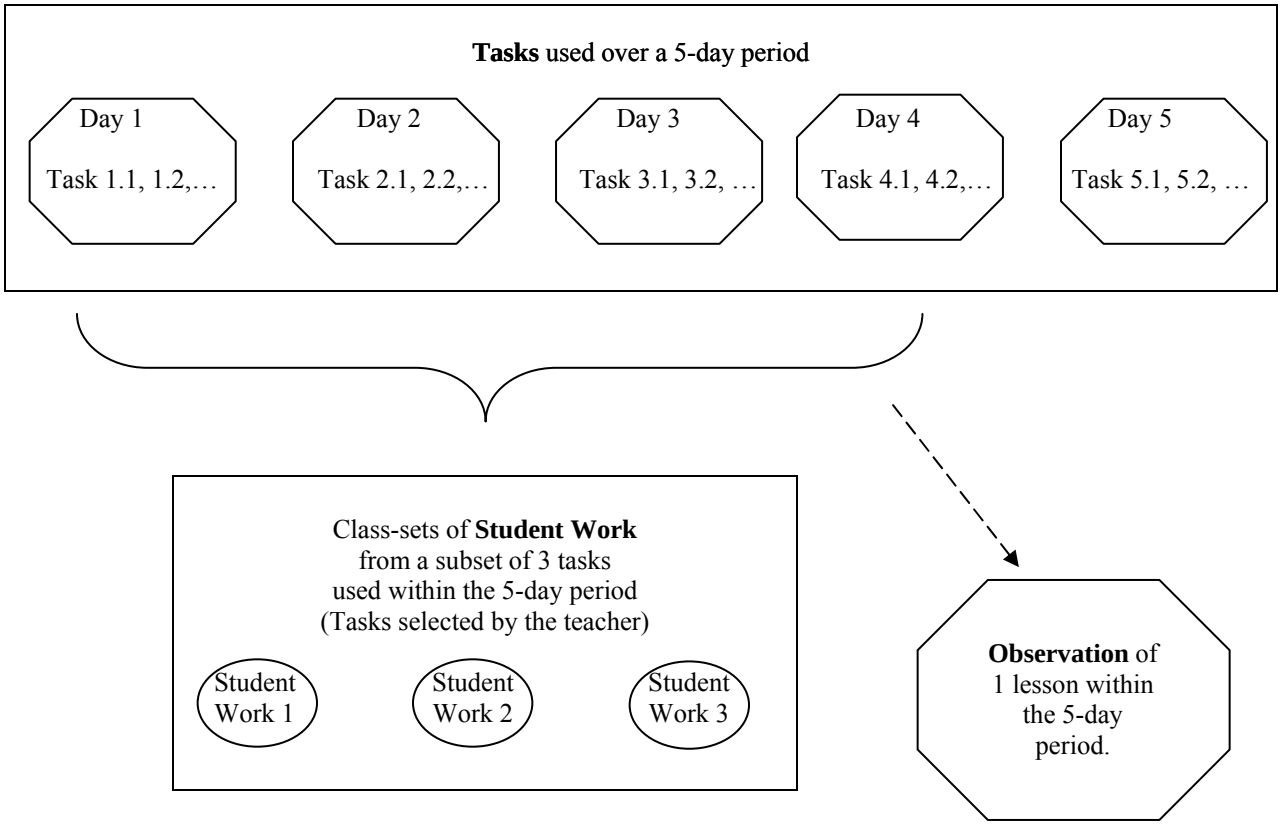


Figure 3.2. Diagram of the data-sets for each data collection.

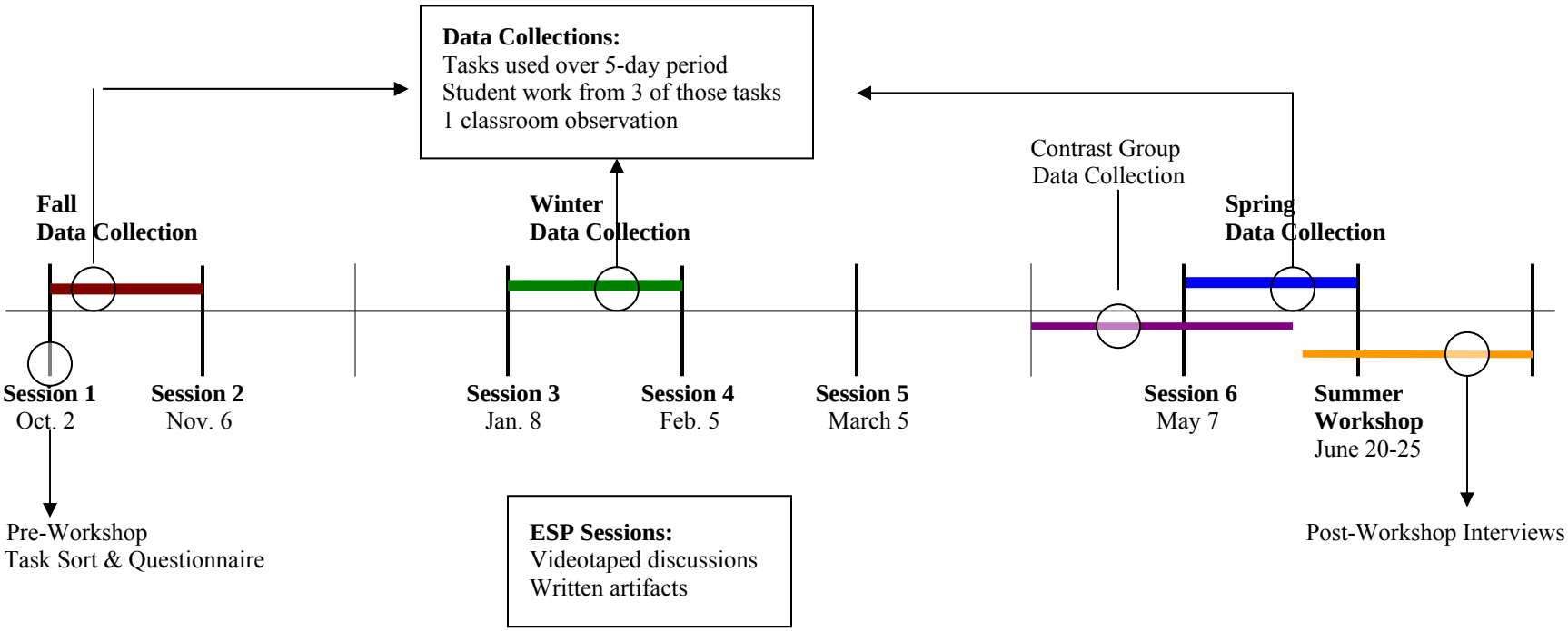


Figure 3.3. ESP data collection timeline.

ESP teachers completed the task sort as an individual written activity during the first ESP session in October 2004 as the pre-test and during the final ESP session in May 2005 as the post-test. Teachers in the contrast group completed the task sort in May 2005. To complete the task sort, participants were provided with 16 cards containing a mathematical task on the front. On the back of the card was a prompt asking participants to: (1) rate the task as High, Low, or Not Sure, and (2) provide a rationale for the rating. Once all tasks have been rated, participants were asked to list their criteria for rating a task as High and Low. The task sort was designed to allow subjects to categorize the tasks based on several possible criteria. The tasks in the task sort differ with respect to the level of cognitive demand and other features such as mathematical content, use of representations (i.e., diagrams or graphs) or manipulatives, use of a context, or requirement of an explanation (Smith, et al., 2004; Smith & Stein, 1998). While many of these features are often associated with high-level mathematical tasks, the task sort was purposefully constructed to contain tasks with similar surface features but different levels of cognitive demand. For example, two tasks that are set in a context (i.e., both are "word problems") or two tasks that contain a prompt to "explain" may differ in their level of cognitive demand. The task sort can be engaged in at some level by all subjects; even those with no recognition of different levels of cognitive demand could devise criteria for categorizing a task as High or Low based on mathematical content or other surface features. The task sort that was used as the pre- and post-test in this investigation is provided in Appendix 3.2.

3.3.2. Collections of Tasks

At three points in the school year (Fall, Winter, and Spring), ESP teachers were asked to collect the mathematical tasks used during instruction over a five-day period for one

mathematics course (hereafter referred to as the *task collection*) (see Figure 3.2). For the purposes of the task collection, tasks were defined to teachers as, "any mathematical problems, exercises, examples, individual or group work, that students encounter from when the bell rings to begin the class period until the bell rings to end the class period." Teachers were provided with individual folders marked Day 1 through Day 5 and a Task Log Sheet for each day. On the Task Log Sheet, teachers were asked to identify the placement of the task within the lesson (i.e., warm-up, main instructional task or concept development, practice, review, homework, etc) and to estimate the amount of time spent on the task. The directions and materials for the task collection were provided to teachers during the first ESP professional development session and again prior to each data collection. The "Directions for Task Collection" and the "Task Log Sheet" are provided in Appendices 3.4 and 3.5, respectively.

The purpose of the task collection is to provide an indicator of the level of cognitive demand of the tasks that teachers use to engage students in learning mathematics over a period of time. An *indicator* is defined by Clare and Aschbacher (2001) as "a statistic that measures outcomes or important dimensions of a system in comparison with a standard over time. The purpose of indicators is to describe the relative functioning of the system and point the way toward improving the system (p. 40)." In this study, the levels of cognitive demand served as the standard used to compare the tasks that teachers use in their classrooms over a period of time. The indicators derived from the task collections described students' opportunities to engage with high-level mathematical tasks in the teachers' classrooms.

3.3.3. Collections of Students' Work

Teachers were asked to collect class-sets of student work for three of the tasks within each task collection. Research conducted by Matsumura and colleagues at the National Center for Research on Evaluation, Standards, and Student Testing (CRESST) (Matsumura, 2003; Matsumura, et al., 2002; Clare & Aschbacher, 2001; Clare, 2000) determined that three sets of student work, consisting of 4 samples each and rated by two raters, yielded a generalizability co-efficient high enough (i.e., $G > .80$) to ascertain the validity of using assignments and student work as indicators of classroom practice. Although the CRESST studies focused on reading comprehension assignments, a recent study by Boston & Wolf (2004, 2006) suggest that mathematics assignments tended to be even more stable within teachers ($G \geq .91$). Based on these findings, in this study, three class-sets of student work were collected in each data collection period (hence, nine class-sets of student work per teacher over the course of the 2004-2005 school year) and were scored by two raters. The design of the student-work collection in the present investigation thus provides a valid indicator of classroom practice in each of the task collection periods.

Teachers were provided with individual folders and a "Student Work Cover Sheet" (Appendix 3.6) marked for each of the three collections of student work. Teachers could submit student work for any three tasks in the data collection, at their own discretion. To preserve anonymity, teachers were reminded to blind students' names before copying the students' work. Teachers were also asked to identify high, medium, and low samples of students work from among the set. In this way, teachers' expectations for students learning and the cognitive processes for which students were actually held accountable can be determined and can serve as indicators of what the teacher values in students' work on the task (Doyle, 1988). Color-coded

stickers were provided to denote the samples of high, medium, and low student work. Procedures for student work collection (Appendix 3.7) were provided to teachers during the first ESP session and prior to each data collection.

3.3.4. Lesson Observations and Lesson Interviews

For ESP teachers, one lesson observation was conducted during each of the data collection periods in the Fall, Winter, and Spring. The purpose of the lesson observations was to provide a basis from which to generalize the classroom practices, with respect to the selection and implementation of high-level mathematical tasks, of individual teachers and of the group of teachers at the beginning, middle, and end of the school year in which they participated in the ESP professional development sessions. The observations lasted an entire class period and were not videotaped or audiotaped. The lesson observer scripted as much of the lesson as possible, specifically attending to the instructional task(s), how the instructional task was presented to students, the interactions that occurred as students worked on instructional tasks, and the exchanges that occurred during any whole-group discussions. Immediately following the lesson, the observer created a timeline of the lesson activities and coded specific features of the lesson (as described in Section 3.4.4). When possible, brief audiotaped interviews were conducted with teachers before the observed lessons to determine teachers' goals and expectations for students' learning and after the lesson to elicit teachers' conceptions of whether and in what ways the goals and expectations were fulfilled. While single classroom observations provide only a snap-shot of teachers' instructional practices at three points in the school year, the lesson observations were compared to other indicators of classroom practice collected within this investigation (i.e.,

tasks and student work). Hence, triangulation of data provided a portrait of the instructional practices of individual teachers and of the group of teachers at different points in time.

3.3.5. Professional Development Observations and Artifacts

Teachers in the ESP Project attended six professional development sessions during the 2004-2005 school year concurrent with this study. The sessions were videotaped and individual and group artifacts (i.e., solutions to tasks, lists created during discussions, written reflections, etc.) were collected. The videotapes and artifacts serve as records of teachers' participation and experiences within the professional development sessions. The presence or absence of changes in teachers' knowledge and instructional practices, within individuals and amongst the group, were compared to their opportunities for learning and their engagement with these opportunities as evident in the videotapes and written artifacts. Conversely, the nature of individual teacher's participation in the professional development sessions were compared to the extent of changes in their knowledge and instructional practices.

3.3.6. Data from the Contrast Group

Teachers in the contrast group completed the task sort instrument and were observed once in May 2005. The task sort provided a measure of the teachers' knowledge of the cognitive demands of mathematical task, and the classroom observation provided a snapshot of the teachers' instructional practices. Data from the contrast group was compared to the Spring data collection from the study group to determine whether the knowledge and instructional practices of teachers who participated in the ESP professional development workshop differed from teachers who did not participate in the workshop.

Figure 3.4 provides a summary of the data collection. In the following section, procedures for coding the data will be described.

3.4. Coding

Data was coded to enable both quantitative and qualitative analyses. All quantitative data was coded by the researcher in this investigation, and a stratified random sample of the pre/post tests, tasks, and student work was coded by a knowledgeable, trained rater to determine reliability. A subset of classroom observations were conducted and coded by two raters, as well. Reliability measures are discussed more thoroughly within each data source.

Quantitative data on the task collections, student work, and lesson observations was obtained using two dimensions of the Instructional Quality Assessment "Academic Rigor in Mathematics" (IQA AR-Math) rubric (Boston & Wolf, 2004, 2006; Junker, Matsumura, Crosson, Wolf, Levison, Weisberg, & Resnick, 2004); *Potential of the Task* and *Implementation of the Task* (Appendix 3.8). Derived from QUASAR research on mathematical tasks (e.g., Stein, et al., 1996; Henningsen & Stein, 1997), the score levels for these two dimensions are based on the cognitive demands of mathematical tasks (*Potential*) and the cognitive processes evident in the lesson or in student's work (*Implementation*). In each dimension, descriptors for score levels 3 and 4 are consistent with characteristics of high-level cognitive demands. Levels 3 and 4 differ with regard to (1) the complexity of the task or of the mathematics in the task and (2) the explicitness of the mathematical connections or reasoning required by the task or made during the task implementation. Score levels 1 and 2 consist of low- level cognitive demands. Level 2 reflects "procedures without connections," and Level 1 is characterized by "memorization" or "no mathematical activity." Hence, an important demarcation line exists between the score levels

Categorization	Focus of Data Collection	Types of Data	Person(s) Collecting the Data	Method	Frequency of Data Collection Pre Mid Post	Research Question Being Addressed
Teachers' Knowledge	Cognitive demands of tasks	Written response	Researcher	Task Sort	1 0 1	Question 1
Indicators of Classroom Practice	Level of Cognitive demands of tasks Level of Implementation of tasks Factors that influenced implementation of tasks	Rubric scores Field notes	Project-based Researchers	Task Collections Student Work Collections Lesson Observations	--5 days each -- 3 3 3 1 1 1	Questions 2 and 3
Professional Development Sessions	Use of high-level tasks Instructional practices (descriptions, insights, reflections, changes)	Videotapes Field notes Self-reports	Project-based Researchers	Professional Development Sessions Observation and artifacts	0 3 3 [6 sessions over the course of the study]	Question 4

Figure 3.4 . Summary of data sources.

of 2 and 3 that separates high- and low- level cognitive demands in each dimension of the AR-Math rubrics.

The remainder of this section describes the coding of quantitative data, the focus of the coding for qualitative data, and the reliability coding specific to each data source.

3.4.1. Task Sort

Teachers' individual written responses to the task sort were coded using a system based on Grover's (1989) approach to scoring teachers' verbal responses to open-ended interview questions. Grover's framework for translating an open-ended response into a numerical score is based on the following considerations:

1. identify the dimensions (i.e., general constructs) that are desired to be assessed in the teacher's response;
2. identify the components that indicate competence in each dimension;
3. assign each component a rating scale, with criteria provided for each score level (i.e., identify what 'anchors' represent competence, sub-competence, or incompetence in each component).

Grover suggests creating a scoring matrix in which the components of each construct can be scored independently, then summed into a total score for the construct and for the entire response. In this investigation, teachers' responses to the task sort were scored based on the following dimensions:

1. Ability to differentiate between tasks with high-level and low-level cognitive demands. For each of the 16 tasks in the task sort, participants' responses were scored based on correct identification of the level of cognitive demand of the task.

2. Ability to identify the characteristics of a task that influence students' opportunities for learning (i.e., characteristics that influence the task's level of cognitive demand). For each of the 16 tasks in the task sort, participants' responses were scored based on whether the rationale identifies elements of the task that influence students' opportunities for learning [i.e., descriptors consistent with the level of cognitive demand of the task (see Figure 2.1)] vs. surface level features of the task (i.e., "uses manipulatives," or "requires explanation") or features inconsistent with the identified level of cognitive demand (i.e., "use of models" as a rationale for rating a task as low-level).
3. Ability to list criteria for high-level tasks and for low-level tasks. The criteria were scored on the participants' identification of the types of tasks associated with high-level cognitive demand (i.e., doing mathematics or procedures with connections) or low-level cognitive demands (i.e., procedures without connections or memorization), descriptors consistent with characteristics of tasks at each level, or surface level features consistent with tasks at each level.

The scoring matrix for the task sort is provided in Appendix 3.10. All 48 task sort responses (19 pre-test, 19 post-test, and 10 contrast subjects) were scored by the primary investigator, and a double-blind, stratified random sample of 10 of the responses (20.8%; 4 pre-test, 4 post-test, 2 contrast group) was scored by a knowledgeable rater (i.e., a mathematics education doctoral student) to check the consistency of the scores and the reliability of the scoring matrix. Agreement of 92.9% was reached on the item-by-item scores of the two raters.

3.4.2. Task Collection

ESP teachers submitted task collection packets from five consecutive days of instruction in the Fall, Winter, and Spring. Teachers were asked to provide all "mathematical problems, exercises, examples, individual or group work, that students encounter from when the bell rings to begin the class period until the bell rings to end the class period;" however, only the tasks determined to be the main instructional tasks for each of the five lessons were coded and analyzed. For the majority of lessons, the main instructional task could be clearly identified by the following criteria: 1) it was described on the teachers' Task Log Sheet as the main instructional task, 2) it fit coherently into the sequence of learning tasks used and the mathematical ideas being developed over the five days of instruction, and 3) it was reported to have taken up the largest amount of instructional time in the lesson. Two dilemmas arose which required a decision as to which task was the main instructional task in the lesson: 1) two tasks equally satisfied all three criteria, or 2) the task that consumed the most instructional time was not the main instructional task in the lesson (i.e., the task was used as a "problem-of-the-day" or a warm-up activity). In the first situation, the scores for the two tasks were averaged into one task score for the lesson. In the second situation, the task that consumed the most instructional time was designated as the main instructional task only if it fit into the sequence of tasks or mathematical ideas in the five lessons; if not, the main instructional task identified by the teacher on the Task Log Sheet was coded.

Eighteen teachers submitted task collection packets in the Fall (18 teachers x 5 main instructional tasks = 90 tasks coded for analysis), 16 teachers submitted task collection packets in the Winter (80 tasks), and 14 teachers submitted task collection packets in the Spring (70 tasks). [Note that 6 teachers did not submit a task collection packet in the Winter and/or Spring. Hence,

task collection data across Fall, Winter and Spring is available for 12 teachers.] The 240 tasks were scored on a scale of 1-4 using the *Potential of the Task* dimension of the IQA AR-Math rubrics (see Appendix 3.8). The *Potential of the Task* dimension assesses the level of cognitive demand that students could potentially engage in by working on the task (i.e., the cognitive demand necessary for students to produce the best possible response to the requirements of the task). The score levels in this dimension are derived from the levels of cognitive demand proposed by Stein and colleagues (1996), as were the levels used to code tasks in the TIMSS 1999 Video Study (USDE-NCES, 2003). To establish the reliability of the rubric, a mathematics education doctoral student received training from the primary investigator and was asked to use the rubric to independently score a set of 32 tasks. Exact agreement was reached on 27 of the 32 tasks (84.4%), and all five disagreements were between the score levels of 3 and 4.

The primary investigator scored all of the main instructional tasks analyzed in this study. Two research specialists on the IQA project independently scored a stratified random sample of 20% of the tasks (48 tasks; 1 task per teacher per data collection) to determine reliability. Exact agreement on the *Potential of the Task* scores between the primary investigator and the IQA research specialists was 83.3% (40/48); 1-point agreement was 93.75% (45/48).

The textbooks used in the teachers' mathematics course for which they were providing the task collection were coded as "innovative" or "traditional." Textbooks coded as "innovative" were identified as "exemplary or promising curricula" by the U.S. Department of Education (USDE, 1999) or were rated highly in the review of mathematics curricula conducted by the American Association for the Advancement of Science (AAAS, 2000). A text not identified by either of those sources was coded as "traditional" unless another source of information strongly indicated otherwise. Other sources of information used to support and verify the coding of the

texts included a review of the text by the researcher in this study and/or a mathematics education doctoral student, materials about the text on the publishers' website, the teachers' descriptions and classifications of their texts, and the publication date.

3.4.3. Student Work

ESP teachers submitted class-sets of student work for three of the tasks in each task collection. Each class-set was scored on a scale of 1-4 using the *Potential of the Task* and the *Implementation* dimensions of the IQA AR-Math rubric (Appendix 3.8). While the *Potential of the Task* dimension assesses the level of cognitive demand in the best possible student-response to the task, the *Implementation* dimension assess the level of cognitive demand over the set of actual student-responses to the task. The score for this dimension is holistic, reflecting the cognitive processes evident in the critical mass of student work samples for a given task (i.e., the score level of the work produced by most of the students). More specifically, individual papers were scored, and the class set was given the modal score of the set (in the event of bimodal scores, the higher score is chosen). Sixteen teachers submitted student work in the Fall, fifteen teachers submitted student work in the Winter, and thirteen in the Spring. [Ten ESP teachers submitted student work for all three data collections.]

The researcher in this investigation scored all 132 sets of student work. Two research specialists on the IQA project independently scored a stratified random sample of 22.7% of the tasks (30 sets of student work; 2 sets per teacher; 10 sets per data collection) to determine reliability. Exact agreement on the *Potential of the Task* scores for the student work was 86.7%; exact agreement on the *Implementation* scores was 93.3%.

3.4.4. Lesson Observations

Eleven ESP teachers and 10 contrast group teachers participated in classroom observations. The lesson observations were coded using the *Potential of the Task* and the *Implementation* dimensions of the IQA AR Math rubric. Each dimension was scored on a scale of 1-4, as described previously. A checklist based on the factors that influence the maintenance or decline of high-level cognitive demands (Henningsen & Stein, 1997) was utilized to guide the scoring of the *Implementation* dimension and serves as a system for obtaining and coding qualitative data from the lesson observations. The lesson checklist and the rubrics for scoring task *Potential* and *Implementation* are provided in Appendices 3.8 and 3.9, respectively. During the observation, the observer scripted the lesson, capturing as much dialogue and detail as possible. Immediately following the lesson, the observer used the field notes to complete the lesson checklist and to rate the *Potential* and *Implementation* of the task.

The primary investigator and a mathematics education doctoral student conducted and rated the lesson observations for this study. Prior to conducting the lesson observations, the two raters observed and independently scored video-taped mathematics lessons to obtain consistency and reliability in using the IQA rubrics. At that point, three pilot lessons were observed by both raters and were then rated independently. Exact agreement of 100% was reached for *Potential* and *Implementation* of the task. Rater calibration was conducted prior to the Winter data collection using the IQA rater-training materials and two additional pilot observations (in which 100% agreement was obtained).

3.4.5. Professionals Development Observations and Artifacts

The six ESP professional development sessions were videotaped, and segments of the videos were earmarked as evidence of teachers' opportunities to consider the cognitive demands of mathematical tasks, the selection or implementation of cognitively challenging mathematical tasks, or the use of cognitively challenging tasks in their own classrooms. Verbal exchanges and written artifacts from the professional development sessions were coded for (1) self-reports of changes in teachers' knowledge or instructional practices, (2) indications of the development of or struggle with new ideas related to the cognitive demands, selection, or implementation of cognitively challenging tasks, or (3) evidence that any of the QUASAR frameworks (Figures 2.1, 2.2, and 2.3) were used to plan, teach, or reflect on an instructional episode (in their own classroom or in a video or narrative case). The selected verbal exchanges were transcribed from videotape.

Following Sessions 2 through 6, ESP teachers were asked to "plan, teach and reflect on a lesson involving a high-level task" (see Figure 3.1) and to bring written reflections or other artifacts from the lesson to the following ESP session. These assignments were analyzed to determine 1) whether teachers had based the lesson on a specific task used within the ESP sessions; and 2) whether the task and/or student work from the lesson was submitted within the teacher's data collection, specifically in the Winter data collection that occurred in the month following Session 3 and in the Spring data collection that occurred in the month following Session 5.

The analyses of all the data sources are presented in the following section.

3.5. Analysis

The analysis of data presented in this section is organized according to data source. For each data source, statistical tests were used where appropriate and feasible to identify statistically significant increases in teachers' knowledge and instructional practices with respect to the selection and implementation of cognitively challenging instructional tasks. The significance level was set at $p < .05$, and one-tailed statistical tests were used when applicable. Qualitative analyses were used to describe 1) the nature of these changes and 2) the professional development experiences of teachers who exhibited significant changes in knowledge and practices versus those who did not. In this way, statistical tests determine whether significant differences in teachers' knowledge and instructional practices have occurred, and qualitative analyses describe what these changes are and how they might be connected to teachers' experiences in the ESP professional development sessions.

3.5.1. Pre- and Post-Workshop Task Sort

Task sorts were coded to yield a numeric score that serves as an indication of teachers' knowledge of the cognitive demands of mathematical tasks. Descriptive statistics were used to assess the pre- and post-workshop task sort scores from ESP teachers and the task sort scores from teachers in the contrast group. ESP teachers' scores on the pre- and post-workshop task sorts were compared using one-tailed Wilcoxon Signed-Rank tests for non-parametric, paired data to determine whether teachers' knowledge of the cognitive demands of mathematical tasks increased following their participation in the ESP professional development sessions. A two-way ANOVA, with "curriculum type" and "time" as the grouping variables, was conducted to determine whether the use of a reform vs. traditional curriculum in teachers' classrooms

influenced their knowledge of the cognitive demands of mathematical tasks before and after their participation in ESP. ESP teachers' and contrast group teachers' task sort scores were compared using Mann-Whitney tests (one-tailed) to analyze differences in the knowledge of cognitive demands of mathematical tasks of the two groups. When statistically significant differences were found, the nature of the differences were analyzed and described qualitatively.

3.5.2. Collection of Tasks

Descriptive statistics were used to assess the level of cognitive demand of the main instructional tasks used in ESP teachers' classrooms for each data collection. Increases in teachers' selection of high-level tasks over time were analyzed in two ways: by comparing differences in the mean task scores between data collections and by comparing the number (and percent) of tasks at each score level for each data collection. A two-way ANOVA (curriculum type by time) was used to identify increases in the mean task scores between data collections and to identify whether teachers' use and/or increased use of high-level tasks was influenced by curriculum type (reform vs. traditional). Mann-Whitney tests and chi-squared tests were then used to supplement the results of the ANOVA. Mann Whitney tests were used to determine when the significant increases in task scores occurred and to analyze the differences in the number of high-level tasks used per teacher in each data collection. Note that an increased number of tasks scoring a 2 vs. 1 or scoring a 4 vs. 3 could be detected by the ANOVA as a significant increase in mean task scores between data collections, but would not represent a true change in teachers' use of high-level tasks. Hence, the number and percent of tasks at each score level were determined for each data collection, as well as percents of tasks rated as having high-level (score of 3 or 4) vs. low-level (score of 1 or 2) cognitive demands. Chi-squared tests were performed to

identify significant differences in the number of high vs. low–level tasks between each of the data collections (3 x 2; Fall/Winter/Spring by H/L).

3.5.3. Collections of Student Work

Each class-set of student work was given one score for *Potential of the Task* and one score for *Implementation* using the IQA AR-Math rubric. Descriptive statistics were used to assess the cognitive demand of the tasks (*Potential*) and the cognitive processes evident in students' written work for solving the tasks (*Implementation*). Mann-Whitney tests were used to identify differences in the mean *Potential* scores between data collections and in the mean *Implementation* scores between data collections. Wilcoxon Signed-Rank tests were used to compare the *Potential* and *Implementation* scores for set of student work within data collections to determine whether the cognitive demands of the student work task declined significantly during implementation. Chi-squared tests were used to identify changes between data collections in the number of student-work tasks that began as high-level (i.e., a score of 3 or 4 for *Potential*) and remained high-level (i.e., a score of 3 or 4 for *Implementation*) during implementation (3 x 2; F/W/S by H/L). Two-way ANOVAs were used to control for curriculum type, as described previously. The *Potential* and *Implementation* scores for student work between data collections were compared qualitatively to identify patterns of maintenance or decline of high-level tasks and whether these patterns changed over time.

3.5.4. Lesson Observations

Lesson observations were scored for the *Potential of the Task* and *Implementation* using the IQA AR-Math rubric. Descriptive statistics were used to assess the cognitive demand of the

lesson tasks (*Potential*) and the cognitive processes in which student engaged during the lesson (*Implementation*) for each data collection and for the lesson observations of teachers in the contrast group. Note that since the lesson observations occurred within the five-day data collection period, the lesson observation tasks are a subset of the tasks previously coded for *Potential* and analyzed with the task collection. Hence, this analysis focuses on changes in the lesson *Implementation* scores (i.e., teachers' ability to maintain high-level cognitive demands during instruction) between data collections and between teachers in the ESP vs. contrast groups. Differences in the mean lesson implementation scores between data collections and between groups of teachers were assessed using Mann-Whitney tests. Due to small sample size, observational comparisons were used to identify changes in the number of lesson tasks that began as high-level (i.e., a score of 3 or 4 for *Potential*) and remained high-level (i.e., a score of 3 or 4 for *Implementation*) between data collections.

Qualitative data on the lesson observations obtained using the IQA Lesson Checklist and the observer's field notes were analyzed to determine the factors that influenced students' engagement with high-level cognitive processes (e.g., Stein, et al., 1996), the patterns of maintenance or decline of high-level cognitive demands (e.g., Henningsen & Stein, 1997), and whether these factors and patterns changed over time. In doing so, the analyses provided a description of ESP teachers' instructional practices that illustrates and supports the findings of the quantitative analyses. Interviews before and after each lesson observation were transcribed to identify teachers' comments regarding the level of cognitive demand of the task or task implementation, expectations for students, and factors that supported or inhibited students' engagement with high-level cognitive processes during the lesson. For selected teachers, lesson narratives were constructed to describe the nature of instruction in their classrooms at different

points in time. Written artifacts or verbal comments from the teachers were used to supplement the lesson narratives when the teacher's comments provided additional insight into their knowledge and/or instructional practices with respect to the selection and implementation of cognitively challenging mathematical tasks.

3.5.5. Professional Development Artifacts and Observations

The written artifacts and videotaped observations from the ESP professional development workshop were analyzed to provide descriptive data on teachers' experiences and participation in the ESP workshop. Instances in teachers' experiences and participation in the ESP workshop that can be reasonably associated with changes in teachers' knowledge or instructional practices were identified and used to provide descriptions of teachers' opportunities to consider the level of cognitive demand of mathematical tasks, the selection and implementation of cognitively challenging mathematical tasks, or the use of cognitively challenging mathematical tasks in teachers' own classrooms. Self-report data, in the form of teachers' statements (transcribed from videotape) or writings during the professional development sessions were utilized to provide instantiations and describe the nature of the changes (or lack thereof) in teachers' knowledge and beliefs throughout their participation in the ESP professional development sessions. Furthermore, data on the frequency and nature of teachers' participation in the ESP workshop was compared to the quantitative analyses and lesson narratives. Based on this comparison, narratives of three teachers were constructed to describe the impact of the professional development sessions on teachers who had statistically significant changes in knowledge and instructional practices as compared to teachers who did not.

Hence, the analyses in the current study coordinated multiple sources of evidence to identify, substantiate, and describe teachers' knowledge and instructional practices with respect to the selection and implementation of cognitively challenging tasks following their participation in the ESP professional development sessions. In Chapter 4, the results of the analyses are presented.

4. CHAPTER 4: RESULTS

The results of the data analysis are reported in this chapter, organized into four sections that correspond to the four research question presented in Chapter 1. Section 4.1 describes teachers' knowledge of cognitive demands of mathematical tasks and how this knowledge changed over time. This includes the results of the pre- and post-workshop task sort, comparisons of the task sort scores of teachers participating in ESP (hereafter referred to as the "ESP group" or "ESP teachers") to the scores of teachers in the contrast group, and qualitative descriptions of the differences in task sort responses over time and between groups. Section 4.2 focuses on ESP teachers' selection of high-level mathematical tasks. This section presents the analyses of the mean task scores and the percentage of tasks at a high vs. low level of cognitive demand from the Fall, Winter, and Spring data collections. Section 4.3 focuses on ESP teachers' implementation of high-level mathematical tasks, and is divided into two sections that address the implementation of tasks in the collection of student work and in the lesson observations, respectively. In the lesson observations, comparisons between the ESP group and the contrast group are presented. Section 4.4 describes the role of the ESP workshops on changes in ESP teachers' knowledge and instructional practices by identifying teachers' opportunities for learning and by identifying key aspects of the professional development experiences for selected teachers.

4.1. Teachers' Knowledge of the Cognitive Demands of Mathematical Tasks

The results presented in this section pertain to Research Question #1:

> Can teachers identify and characterize mathematical tasks with high-level cognitive demands and mathematical tasks with low-level cognitive demands, and does this change after participation in professional development specifically focused on the selection and implementation of cognitively challenging mathematical tasks?

To answer this question, comparisons were made between ESP teachers' pre- and post-workshop task sort scores and between the task sort scores of ESP teachers and contrast group teachers. The results of these comparisons are presented in the remainder of this section.

4.1.1. Pre- and Post-Workshop Task Sort

The pre- and post-workshop task sort scores serve as an indicator of ESP teachers' knowledge of cognitive demands prior to and following their participation in the ESP workshops, respectively. Nineteen teachers participated in the pre- and post-workshop task sort. For each of the 16 tasks in the task sort, teachers received 1 point for correctly classifying the task as high-level or low-level according to the TAG (see Figure 2.1) (i.e., "doing mathematics" and "procedures with connections" tasks would be classified as high-level; "procedures without connections" and "memorization" tasks would be classified as low-level) and 1 point per task for providing a rationale that identified task features consistent with the task's level of cognitive demand. Teachers also received 0 to 3 points for providing overall criteria for high-level tasks and 0 to 3 points for providing overall criteria for low-level tasks, making the highest possible score on the task sort 38 points. The scoring rubric for the task sort is provided in Appendix 3.10.

Scores on the pre-workshop task sort ranged from 13 to 32 points, with a mean score of 24.21. Post-workshop task sort scores ranged from 19 to 37 points, with a mean score of 28.74. Results of the Wilcoxon Signed-Rank tests for non-parametric, paired data indicate that the increase of 4.53 between the means of the pre- and post-workshop task sorts was significant ($z = 3.15$; $p < .001$ [one-tailed]). These results suggest that ESP teachers' knowledge of the cognitive demands of mathematical tasks increased following their participation in the ESP workshops. Table 4.1 provides descriptive statistics on ESP teachers' task sort scores overall and grouped by curriculum type.

Table 4.1.
Descriptive Statistics on Task Sort Scores

		Pre-Workshop	Post-Workshop
	n	Mean (SD)	Mean (SD)
All teachers	19	24.21 (5.75)	28.74 (4.12)
Teachers Using Reform Curricula	10	24.10 (5.78)	28.00 (5.08)
Teachers Using Traditional Curricula	9	24.33 (6.06)	29.56 (2.79)

ESP teachers' knowledge of cognitive demands of mathematical tasks, and the change in this knowledge over time, was not influenced by the use of a reform vs. traditional curricula. As identified in Table 4.1, the pre- and post-workshop task sort scores for the subset of 9 ESP teachers using traditional curricula were slightly higher than the scores for the subset of 10

teachers using reform curricula. Teachers using traditional curricula also increased their task sort scores from pre- to post-workshop by a greater amount than the teachers using reform curricula (i.e., 5.23 as compared to 3.90, respectively). The two-way ANOVA, however, indicated that the difference in means between teachers using reform vs. traditional curricula was not significant ($F = 0.192$; $p = .667$) and that teachers using one type of curricula did not increase in scores significantly more than teachers using the other type of curricula ($F = 0.324$; $p = .577$). ANOVA results also confirm the increases in task sort scores over time ($F = 15.424$; $p < .001$). This dispels the assumption that teachers using reform curricula would have greater knowledge of the cognitive demands of mathematical tasks due to their exposure to a greater number of higher-level tasks in their curricula.

Hence, the increase in ESP teachers' knowledge of the cognitive demands of mathematical tasks following their participation in the ESP professional development workshops almost certainly cannot be attributed to chance or to the type of curricula used in their classrooms. The nature of these increases will be described later in this section, and the analysis of the ESP professional development workshops in Section 4.4 will identify events that might have provided opportunities for this learning to occur.

4.1.2. Comparing ESP teachers to the contrast group

ESP teachers' pre-and post-workshop task sort scores were compared to the task sort scores of the contrast group. The results indicate whether ESP teachers' had a greater knowledge of the cognitive demands of mathematical tasks at the close of the school year than a group of secondary mathematics teachers who did not participate in the ESP workshop.

The task sort scores from the 10 contrast group teachers ranged from 8 to 26 points, with a mean of 17.6. Results of the Mann-Whitney tests comparing the task sort scores of the ESP teachers and the contrast group are listed in Table 4.2. The pre-workshop task sort scores of the 19 ESP teachers were significantly higher than those of the contrast group ($z = 2.32$; $p = .01$ [one-tailed]), indicating that, even prior to their participation in the ESP workshop, ESP teachers had greater knowledge of the cognitive demands of mathematical tasks than the contrast group. However, five ESP teachers had previous exposure to the task sort instrument through their participation in other professional development experiences in the region (personal communication, 10/8/04), and their pre-workshop task sort scores were deleted from the comparison. The scores of the remaining 14 ESP teachers were not significantly higher than the scores of the contrast group ($z = 1.55$; $p = .06$ [one-tailed]). Hence, the 14 ESP teachers with no prior exposure to the task sort at the beginning of the school year can be assumed to have similar knowledge of the cognitive demands of mathematical tasks as the contrast group at the close of the school year. This finding indicates that mere exposure to tasks, curriculum, and teaching throughout the course of a school year does not enable teachers to improve their ability to identify the features of tasks with high- and low-level cognitive demands. In contrast, intervention in the form of professional development experiences specifically focused on the cognitive demands of mathematical tasks appear to provide teachers with such knowledge. The post-workshop task sort scores of all 19 ESP teachers were significantly greater than the task sort scores of the contrast group ($z = 3.95$; $p < .001$ [one-tailed]); in addition, the scores of the subset of 14 ESP teachers with no prior exposure to the task sort also were also significantly greater than the contrast group ($z = 3.63$; $p < .001$ [one-tailed]). This difference indicates that at the close of the school year, the ESP teachers' knowledge of cognitive demands of mathematical tasks

following their participation in the ESP workshop was significantly higher than a group of teachers who had not participated in the ESP workshop.

4.1.3. Descriptive Data on Teachers' Task Sort Responses

Analyses of ESP teachers' pre- and post-workshop task sort scores examined the nature of the increases in teachers' knowledge of the cognitive demands of mathematical tasks; specifically, whether gains in the pre- to post-workshop task sort scores could be attributed to an improvement in teachers' ability to identify high-level tasks, teachers' ability to identify low-level tasks, and/or teachers' ability to describe the features of high- and low-level tasks. Table 4.3 provides data to illustrate the nature of changes in teachers' task sort responses over time.

Teachers' ability to correctly identify high-level tasks did not improve from pre- to post-workshop. Teachers were successful at classifying "Doing Mathematics" (DM) tasks as high-level on the pre-workshop task sort, and no marked improvements were noted on the post-workshop task sort. Of the 95 instances in which DM tasks were classified on the task sort (i.e., 5 DM tasks per teacher times 19 teachers), only 16 incorrect classifications (17%) occurred on the pre-workshop task sort and 15 (16%) occurred on the post-workshop task sort. As shown in Table 4.3, ten teachers incorrectly classified at least one DM task on the pre-workshop task sort and nine teachers did so on the post-workshop task sort.

Table 4.2.
Comparison of Task Sort Scores of ESP Teachers and Contrast Group

	Mean (SD)	Mean Difference vs. Contrast Group
Contrast Group (n = 10)	17.60 (6.13)	NA
Pre- Workshop:		
All ESP teachers (n = 19)	24.21 (5.75)	6.61*
ESP teachers with no prior exposure (n = 14)	22.86 (5.99)	5.26
Post-Workshop		
All ESP teachers (n = 19)	28.74 (5.84)	11.14*
ESP teachers with no prior exposure (n = 14)	29.00 (5.25)	11.40*

*Results are significant at $p < .01$ [one-tailed].

Conversely, teachers had difficulty categorizing "Procedures with Connections" tasks (PWC) as high-level tasks on both the pre- and post-workshop task sort. PWC tasks were the most frequently missed category of tasks, categorized incorrectly as low-level tasks in 52% (49 of 95 instances) of the occurrences of PWC tasks on the pre-workshop task sort and in 49% (47 of 95 instances) of the occurrences of PWC tasks on the post-workshop task sort. Though the percentage of incorrect categorizations decreased slightly, 17 teachers incorrectly classified at least one PWC task on the post-workshop task sort compared to 15 teachers doing so at pre-workshop. PWC tasks were categorized incorrectly three times as often as DM tasks on both the

pre- and post-workshop task sort. Hence, the significant increase in teachers' task sort scores cannot be attributed to an increased ability to identify high-level tasks. Teachers entered the project able to consistently identify DM tasks as having high-level cognitive demands, and teachers did not improve their ability to recognize PWC tasks as having high-level cognitive demands.

Table 4.3 also illustrates that ESP teachers were proficient in identifying low-level tasks on the pre-workshop task sort, and this ability improved slightly over time. On the 76 occurrences in which "Procedures without Connections" (PWOC) tasks were classified (i.e., 4 PWOC tasks per teacher times 19 teachers), PWOC tasks were incorrectly classified as high-level tasks on 20% (15 of 76) of pre-workshop task sort responses and 9% (7 of 76) of post-workshop task sort responses. On the pre-workshop task sort, 15 teachers incorrectly classified at least one PWOC task as high-level, and this number decreased to 7 on at post-workshop. Two "Memorization" (MEM) tasks were on the task sort, creating 38 instances (2 tasks times 19 teachers) where MEM tasks were classified. Of these instances, 11% (4 of 38) were classified incorrectly as high-level on the pre-workshop task sort and none were classified incorrectly on the post-workshop task sort. Four teachers incorrectly classified at least one MEM task as high-level on the pre-test, and no teachers did so at post-test. These results suggest that ESP teachers exhibited a slightly enhanced ability to identify tasks with low-level cognitive demands following their participation in the ESP workshop.

However, the majority of gains in task sort scores over time can be attributed to improvements in teachers' ability to provide appropriate rationales and criteria for high- and low-level tasks. On the pre-workshop task sort, 6 teachers listed criteria inconsistent with features of high- or low-level tasks (i.e., a task is low-level tasks if it contains a diagram; a task is

high-level if it is beyond students' reach) compared to no teachers providing inconsistent criteria on the post-test. On the post-test, all 19 teachers identified characteristics of DM tasks (i.e., open-ended, problem-solving, or specific use of "doing mathematics") in their criteria for high-level tasks. Ten teachers included criteria consistent with PWC tasks, with 9 teachers using the specific terminology, "procedures with connections." The number of teachers who identified "procedures without connections" or something synonymous (i.e., computation, basic skills, drill problems, etc.) in their rationale for low-level tasks increased from 9 to 19, and the number of teachers who listed "memorization" as a feature of low-level tasks increased from 10 to 19. Interestingly, the only category not identified by all 19 teachers on the post-workshop task sort was PWC; though ten teachers included "making connections" as a characteristic of high-level cognitive demands, ESP teachers overall did not improve in their ability to identify a PWC task.

Results of qualitative comparisons also illuminated interesting similarities and differences between the task sort responses of ESP teachers and contrast group teachers. Table 4.4 provides the results of this comparison. Similar to ESP teachers, contrast group teachers experienced the most difficulty with identifying and describing the characteristics of PWC tasks. Five contrast teachers (50%) identified criteria inconsistent with characteristics of high- or low-level tasks, compared to 6 pre-workshop ESP teachers (32%) and 0 post-workshop ESP teachers. Another difference was that ESP teachers' were more likely to use: a) the specific levels of cognitive demand of mathematical tasks from the Task Analysis Guide featured in Figure 2.1 (Stein, et al., 1996); and b) terminology frequently used within the ESP workshops to describe key features of high-level and low-level tasks (i.e., representations, generalizations, connections, procedural).

Table 4.3.
Analysis of the Task Sort Responses by Level of Cognitive Demand (Stein, et al., 1996) (n = 19 teachers)

Level of Cognitive Demand	# of Tasks	Total # of classifications[a]	# of incorrect classifications		# of teachers incorrectly classifying a task at that level		# of teachers identifying the category in their criteria	
			Pre-Workshop	Post-Workshop	Pre-Workshop	Post-Workshop	Pre-Workshop	Post-Workshop
High Level:								
Doing Mathematics	5	95	16	15	10	9	14	19
Procedures with Connections	5	95	49	47	15	17	5	10
Low-Level:								
Procedures without Connections	4	76	15	7	12	7	9	19
Memorization	2	38	4	0	4	0	10	19

[a]Total number of classifications is determined by multiplying the number of tasks at that level by 19 (the number of teachers)

The qualitative findings help to substantiate that the increases in ESP teachers' task sort scores were not the effect of the repeated measures design (i.e., the scores did not improve simply because teachers were completing the task sort for the second time), nor of teachers learning the "correct answers" to the task sort. Increases in task sort scores can be attributed to an improvement in teachers' ability to identify and characterize tasks with low-level cognitive demands and to provide overall criteria for high- and low-level tasks. Differences in ESP teachers' task sort responses from pre- to post-workshop and between the ESP and contrast groups indicate that the ESP teachers learned to identify and describe tasks with high and low levels of cognitive demand using characteristics from the Task Analysis Guide and other ideas made salient in the ESP workshop.

Table 4.4.
Qualitative Comparison of Task Sort Responses between ESP Teachers and Contrast Group Teachers

	ESP Teachers (post-workshop)		Contrast teachers	
Use of specific Level of Cognitive Demand:				
Once	19	(100%)	3	(30%)
Twice	15	(79%)	0	(0%)
3+	9	(47%)	0	(0%)
Use of terminology frequently used in ESP:				
Representations	9	(47%)	3	(30%)
Multiple solution methods; open-ended	15	(79%)	1	(10%)
Generalization; generalize	9	(47%)	0	(0%)
Making connections	9	(47%)	0	(0%)
Procedural; computation; etc.	16	(84%)	4	(40%)

4.2. Teachers' Selection of Instructional Tasks

This section addresses research question #2:

> Do teachers use mathematical tasks with high-level cognitive demands to engage students in learning mathematics, and does this change during and following participation in professional development specifically focused on the selection and implementation of cognitively challenging mathematical tasks?

Collections of tasks were analyzed by comparing the mean task scores between data collections and by comparing teachers' use of high- vs. low-level instructional tasks between data collections. Findings from both types of analyses, respectively, will be discussed in this section. [Note that contrast group teachers did not submit collections of tasks; hence, in this section, "teachers" refers to ESP teachers exclusively.]

4.2.1. Differences in Mean Scores between Task Collections

The five main instructional tasks in project teachers' data collections were scored on a scale of 1-4 using the *Potential of the Task* dimension of the IQA AR-Math rubrics (Boston & Wolf, 2004, 2006). The mean of a teacher's five scores serves as an indicator of the level of cognitive demand of the tasks used in the teacher's classroom over the 5-day data collection. At least one task collection exists for 18 teachers, though only 12 teachers provided tasks in all three data collections (see Appendix 4.1 for a discussion of attrition). Descriptive statistics are reported in Table 4.5 for all of the data available at each point in time and for the subset of 12 teachers with complete data for Fall, Winter, and Spring. A comparison of the task means and confidence intervals indicates that 1) the mean task scores from the subset of 12 teachers with complete data sets is representative of the entire data set; and 2) the mean task scores increased

over time. Two types of tests were performed to identify whether these increase were significant and represented true increases in teachers' use of high-level tasks.

A two-way ANOVA test was used to identify whether teachers' task scores increased over time and whether task scores (and the increases over time) were influenced by the type of curriculum (reform vs. traditional) used in the teachers' classroom. The ANOVA was conducted on the subset of 12 teachers for which data was available for all three data collections. Table 4.6 provides descriptive statistics for task scores grouped by curriculum type. Instructional tasks used by teachers with reform curricula were higher at each data collection, with an approximate difference of 0.5 in Fall and Winter and 0.2 in Spring. However, results of the ANOVA indicate that these differences were not significant ($F = 3.61$; $p = .09$). Curriculum type did not influence the level of tasks used in ESP teachers' classrooms. An insignificant interaction between time and curriculum ($F = 1.12$; $p = .35$) indicates that the increase in task levels over time was not significantly greater in one group than in the other; teachers using each type of curricula experienced similar gains in the levels of the tasks used in their classrooms. These gains were significant ($F = 7.35$; $p < .01$), indicating that time had a significant effect on teachers' use of higher-level tasks. Hence, ESP teachers significantly increased the level of tasks used in their classrooms throughout their participation in the ESP workshop, and these gains were not influenced by curriculum type.

To supplement the results of the ANOVA, Mann-Whitney tests were used to compare the differences in mean task scores between data collections, using all data available. Mean task scores for all available data (see Table 4.5) increased from 2.54 in Fall, to 2.93 in Winter (an increase of 0.39), to 3.01 in Spring (an increase of 0.08 from Winter and 0.47 from Fall). The results of the Mann-Whitney tests indicate that a significant increase in mean task scores

Table 4.5.
Descriptive statistics on the Potential of the Task Scores for the Task Collection

	Number of Teachers	Mean (SD)	95% Confidence interval
Fall			
All data	18	2.54 (0.48)	2.30 - 2.77
Subset	12	2.59 (0.52)	2.28 - 2.89
Winter			
All data	16	2.93 (0.55)*	2.64 - 3.22
Subset	12	2.90 (0.47)	2.65 - 3.15
Spring			
All data	14	3.01 (0.53)*	2.70 - 3.31
Subset	12	3.09 (0.49)*	2.77 - 3.41

*Significant increase from Fall at $p < .05$

Table 4.6.
Descriptive Statistics on Potential of the Task Scores Grouped by Curriculum Type

	Mean (SD)		
Data Collection	Reform (n = 6)	Traditional (n = 6)	Difference in Means
Fall	2.83 (0.43)	2.35 (0.53)	0.48
Winter	3.18 (0.27)	2.62 (0.48)	0.56
Spring	3.18 (0.60)	2.99 (0.37)	0.19

occurred between Fall and Winter ($z = 1.79$; $p = .04$ [one-tailed]) and between Fall and Spring ($z = 2.34$; $p < .01$ [one-tailed]), while no significant increase occurred between Winter and Spring ($z = 0.25$; $p = .40$ [one-tailed]).

Results of the Mann-Whitney tests and the ANOVA suggest that the level of cognitive demand of the tasks used in ESP teachers' classroom increased significantly from the beginning to the end of teachers' participation in the ESP workshop, in ways that are not likely the result ofchance nor influenced by the teachers' curricula. Most of this increase occurred between the Fall and Winter data collections, following teachers' participation in the first three ESP sessions (October, November, and January). Task levels continued to increase slightly between Winter and Spring over the course of teachers' participation in the remaining three ESP sessions (February, March, and May), resulting in a significant increase between the beginning and end of teachers' participation in ESP. In Section 4.4, the events of the ESP workshop that might have contributed to teachers' use of higher level tasks in their classrooms between the Fall and Winter data collections, and the events in the last three ESP sessions that may have continued to enhance these gains will be discussed.

4.2.2. Differences in the Percent of High-Level Tasks between Data Collections

The next set of tests determined whether the significant increases in the mean task scores over time were truly indicative of teachers' increased use of high-level tasks. A myriad of increases in task scores could affect the mean but not the percent of tasks at a high vs. low level of cognitive demand (i.e., increases between score levels 1 and 2 or between score levels 3 and 4). The number and percent of tasks at each score level is portrayed in Table 4.7, for all available data. (Note that in Table 4.7, n is the number of teachers submitting tasks in each data collection,

and the number of tasks is determined by multiplying n x 5 [the number of teachers times 5 main instructional task each]). Chi-squared tests compared the number of high-level (i.e., score of 3 or 4) and low–level (i.e., score of 1 or 2) tasks in each data collection (3 x 2; Fall/Winter/Spring by H/L), using all of the data available at each time period and using only the 12 teachers with a data set for Fall, Winter, and Spring to assess whether teachers increased their use of high-level tasks and to determine whether the tasks used by teachers with complete data collections were somehow different than the tasks used by teachers with incomplete data collections. Results indicate that the number of high-level instructional tasks used in teachers' classrooms increased significantly over time in a way that could not be attributed to chance, for all available data ($\chi^2(2) = 16.18$; $p < .01$) and for the subset of 12 teachers with complete data sets ($\chi^2(2) = 13.72$; $p < .01$).

Table 4.7.
Number (and Percent) of Tasks at each Score Level for Potential of the Task

		Low-Level Cognitive Demands			High-Level Cognitive Demands		
	# of Tasks	Score = 1	Score = 2	**Total L-L**	Score = 3	Score = 4	**Total H-L**
Fall							
n = 18	90	3 (3%)	47 (52%)	**50 (56%)**	31 (34%)	9 (10%)	**40 (44%)**
n = 12	60	3 (5%)	29 (48%)	**32 (53%)**	20 (33%)	8 (13%)	**28 (47%)**
Winter							
n = 16	80	3 (4%)	23 (29%)	**26 (33%)**	33 (41%)	21 (26%)	**54 (67%)**
n = 12	60	3 (5%)	16 (27%)	**19 (32%)**	27 (45%)	14 (23%)	**41 (68%)**
Spring							
n = 14	70	0 (0%)	19 (27%)	**19 (27%)**	33 (47%)	18 (26%)	**51 (73%)**
n = 12	60	0 (0%)	13 (22%)	**13 (22%)**	42 (48%)	18 (30%)	**47 (78%)**

Due to the nature of the IQA AR-Math rubric, changes in score levels reflect qualitative differences in teachers' instructional tasks over time. In the Fall data collection, over half (52%) of the main instructional tasks were categorized as "procedures without connections," such as performing computations or following a set of rote, prescribed steps in an algorithm. Only 10% of tasks provided opportunities for students to explicitly demonstrate mathematical reasoning and understanding, such as having to demonstrate connections between mathematical representations or describe the reasoning behind a mathematical procedure. Throughout the course of teachers' participation in the ESP workshop, teachers began to incorporate a greater percentage of tasks with high-level cognitive demands. More than half of the main instructional tasks in the Winter and Spring data collections (67.5% and 72.85%, respectively) provided students with opportunities to engage in high-level thinking and reasoning, whether implicit (i.e., a score or 3) or explicit (i.e., a score of 4) in the task demands.

The average number of high-level tasks per teacher (using all available data) increased from 2.22 in Fall to 3.38 in Winter, an increase of 1.16 high-level tasks per teacher between Fall and Winter. The number of high-level tasks per teacher continued to increase to 3.64 in Spring, resulting in an increase of 1.42 overall. The significance of the increases was assessed using Mann-Whitney tests. The results reflect the same pattern of significance reported on the previous Mann-Whitney tests comparing the differences in the mean task scores – significant increases from Fall to Winter ($z = 2.23$; $p = .013$ [one-tailed]) and from Fall to Spring ($z = 2.33$; $p < .01$ [one-tailed]), and no significant increase from Winter to Spring ($z = .29$; $p = .295$ [one-tailed]).

Hence, increases in the mean level of cognitive demand were indicative of the increased use of high-level instructional tasks in ESP teachers' classrooms. Significant increases in the mean task scores, the percent of tasks at a high- vs. low level of cognitive demand, and in the

number of high-level tasks per teacher occurred between Fall and Winter and between Fall and Spring. Analysis of the professional development sessions will identify events that catalyzed these changes between Fall and Winter and served to sustain them at slightly higher levels between Winter and Spring.

4.3. Teachers' Implementation of Tasks

This section addresses Research Question #3:

> Do teachers implement mathematical instructional tasks in ways that support students' engagement with high-level cognitive demands, and does this change during and following participation in professional development specifically focused on the selection and implementation of cognitively challenging mathematical tasks?

This question will be answered by examining the results of the analyses of the student work and lesson observations.

4.3.1. Collections of Student Work

Teachers submitted three class-sets of student work in each data collection. Sets of student work were scored for the level of cognitive demand of the tasks (*Potential*) and the cognitive processes evident in students' written work for solving the tasks (*Implementation*) using the IQA AR-Math rubric. Sixteen teachers provided student work in the Fall, fifteen teachers provided student work in the Winter, and thirteen teachers provided student work in the Spring (Appendix 4.1 provides a discussion of attrition). Descriptive statistics on all of the available data are provided in Table 4.8. Data listed under the heading "Student Work Means

(SD)" in Table 4.8 identifies the student work *Potential*[2] and *Implementation* means for each data collection. Based on this data, student work *Implementation* scores were analyzed for increases in the means between data collections, as an indicator of whether teachers' were able to implement instructional tasks at a high-level of cognitive demand. Data under the heading "Number of SW Tasks rated as High-Level" in Table 4.8 identifies the number of student work tasks that were rated as high-level for *Potential* and for *Implementation.* These data were analyzed for increases in the number of high-level implementations between data collections, to determine whether increases in the mean *Implementation* scores reflect actual increases in the number of high-level implementations (rather than increases from a score of 1 to 2 or from a score of 3 to 4 that would affect the *Implementation* mean score but not the number of high vs. low level implementations).

4.3.1.1. Analyzing the Differences in Student Work *Implementation* Means

Mann-Whitney tests were used to determine the significance of the differences in student work *Implementation* means between data collections (see Table 4.8). Between Fall and Winter, the increase in *Implementation* scores of 0.32 was just below the significance level ($z = 1.64$; $p = 0.05$). Student work *Implementation* means continued to increase between Winter and Spring, and though this increase was not significant, it effected an overall increase of 0.59 between Fall and Spring that was significant ($z = 2.94$; $p = 0.002$ [one-tailed]). Hence, teachers consistently increased their level of task implementation throughout the course of their participation in the

[2] In each data collection, the mean *Potential* scores for student work tasks fall within the 95% confidence intervals of the mean *Potential* scores for the collection of instructional tasks (presented in Table 4.5); hence, the three instructional tasks for which teachers submitted student work were representative of the five main instructional tasks in their data collection.

ESP workshops and, by the conclusion of the workshops, were able to implement instructional tasks at a significantly higher level of cognitive demand.

Table 4.8.
Descriptive Statistics on Student Work (SW) scores for Potential and Implementation (for all available data)

		Student Work Mean (SD)		Number of SW tasks rated as High-Level	
	# of Tasks	*Potential*	*Implementation*	*Potential*	*Implementation*
Fall (n = 16)	48	2.63 (0.79)	2.27 (0.57)	23 (48%)	12 (25%)
Winter (n = 15)	45	3.02 (0.81)	2.59 (0.72)	32 (71%)	24 (53%)
Spring (n = 13)	39	3.03 (0.80)	2.86 (0.81)	30 (77%)	26 (67%)

Similar to the argument stated previously regarding increases in task selection means, significant increases in *Implementation* means do not guarantee that teachers improved their ability to implement tasks at a high-level (i.e., at a score of 3 or 4) if the increases occurred between score levels 1 and 2 (implementations that improved but remained low-level) or between score levels 3 and 4 (improvements in implementations that were already at a high-level). A chi-squared test was conducted to compare the number of high- vs. low-level implementations between data collections (see Table 4.8 for the number of tasks implemented at

a high level in each data collection). Results indicate that the number of high-level implementations evident in students' work increased significantly over time in a way that could not be attributed to chance ($\chi^2(2) = 16.11$; $p < .001$). This result also confirms that the significant increases in student work implementation means reflect a true increase in teachers' implementation of tasks at a high-level of cognitive demand.

Implementation scores improved significantly from Fall to Spring, but earlier results indicated that teachers were also using high-level instructional tasks more frequently in Spring. An argument could be waged that using a greater number of high-level tasks enabled teachers to enact a greater number of high-level implementations. This raises the question of whether the increase in task *Implementation* scores is simply a natural consequence of the increase in task *Potential* scores rather than a true reflection of improvements in teachers' ability to maintain high-level cognitive demand during instruction. Another way to analyze improvements in teachers' implementation is to examine the relationship between task *Potential* and task *Implementation* within each data collection. Comparisons between the *Potential* and *Implementation* data listed in Table 4.8 indicate that, in all three data collections, 1) task *Implementation* means are lower than task *Potential* means, and 2) the number of high-level implementations is lower than the number of high-level tasks. These differences suggest that even when teachers selected high-level instructional tasks for use in their classrooms, the high-level tasks were often enacted during the lesson in ways that did not maintain students' opportunities to engage with high-level thinking and reasoning. However, the data in Table 4.8 also indicate that the number of tasks maintained at a high-level during implementation increased from 25% (12 out of 48) in Fall to 67% (26 out of 39) in Spring. This equates to less than 1 high-level implementation per teacher in Fall (i.e., 12 high-level implementations per 16 teachers) as

compared to 2 high-level implementation per teacher in Spring (i.e., 26 high-level implementations per 13 teachers). These statistics suggest that fewer high-level tasks were declining during implementation in Spring than in Fall. A chi-squared test was used to assess whether teachers' ability to maintain the cognitive demands of high-level student work tasks improved over time. For student work tasks that were coded as high-level for *Potential,* Table 4.9 reports the number and percent of implementations that were coded as high-level vs. low-level in Fall, Winter, and Spring. The results of the chi-squared test indicate that significant changes did occur between data collections in the number of student-work tasks that began as high-level (i.e., a score of 3 or 4 for *Potential*) and remained high-level during implementation ($\chi^2(2) = 7.96$; $p = 0.02$). This implies that, throughout their participation in the ESP workshops, teachers improved their ability to maintain high-level cognitive demands as evident in student work.

Qualitative analyses confirm the results of the statistical tests, indicating that fewer high-level tasks declined into "procedures without connections" (i.e., score level 2) over time; subsequently, engagement with high-level cognitive demands was evident in a greater number of students' work over time. Only three student-work tasks increased their score from a 3 for *Potential* to a 4 for *Implementation*; no student-work tasks coded as low-level (i.e., a score of 1 or 2) for *Potential* were subsequently coded as high-level for *Implementation.*

4.3.1.2. Analyzing the Influence of Curriculum on Teachers' Student–Work *Implementation* Scores

Were the changes in student work *Implementation* scores influenced by the use of reform vs. traditional curriculum in teachers' classrooms? To investigate this question, a two-way ANOVA was conducted using the subset of ten teachers who submitted student work in all three data

collections. Table 4.10 provides descriptive data for the student work *Implementation* scores grouped by curriculum type. In each data collection, teachers using reform curricula had higher student work *Implementation* scores than teachers using traditional curricula. Results of the ANOVA test used to determine whether this difference was significant indicated that the main effect for curriculum type ($F = 2.71$; $p = 0.152$) and the interaction between time and curriculum ($F = 0.72$; $p = 0.50$) were both non-significant. Only the main effect for time was significant ($F = 7.95$; $p = 0.004$). Therefore, teachers using reform curricula did not implement student work tasks at a significantly higher level than teachers using traditional curricula, and the improvement in student work *Implementation* scores was not significantly greater in one group than in the other. Similar to findings on teachers' selection of tasks, the type of curriculum did not greatly influence teachers' implementation of student-work tasks.

Table 4.9.
A Comparison of Implementation Scores for Student Work Tasks rated as High-Level for Potential

Data Collection	Number of tasks coded as high-level for *Potential*	Number (%) of tasks coded for *Implementation* as:	
		High-Level	Low-Level
Fall	23	12 (52%)	11 (48%)
Winter	32	24 (75%)	8 (25%)
Spring	30	26 (87%)	4 (13%)

Table 4.10.
Descriptive Statistics on Student Work Implementation Scores Grouped by Curriculum Type

	Curriculum Type	n	Mean	SD
Fall		10	2.40	0.38
	Reform	5	2.67	0.33
	Traditional	5	2.13	0.18
Winter		10	2.60	0.52
	Reform	5	2.67	0.53
	Traditional	5	2.53	0.56
Spring		10	3.05	0.47
	Reform	5	3.20	0.38
	Traditional	5	2.90	0.55

Hence, significant increases in student work *Implementation* means and in the percentage of high-level student work implementations indicate that ESP teachers improved their ability to maintain high-level cognitive demands during implementation.

4.3.2. Lesson Observations

Lesson observations were conducted for eleven ESP teachers and 10 contrast group teachers. ESP teachers were observed once in each data collection and contrast group teachers were observed once in the Spring. Descriptive statistics on the cognitive demand of the lesson

observation tasks (*Potential*) and the cognitive processes in which student engaged during the lesson (*Implementation*) are presented in Table 4.11.

Table 4.11.
Descriptive Statistics on Lesson Observations

	Potential Mean (SD)	*Implementation* Mean (SD)
ESP Teachers (n = 11)		
Fall	2.68 (0.72)	2.45 (0.69)
Winter	3.23 (0.88)	2.86 (0.90)
Spring	3.18 (0.75)	2.91 (0.83)
Contrast Group (n = 10)	2.40 (0.52)	2.20 (0.42)

Table 4.12 provides the differences in mean scores between data collections and between ESP teachers and the contrast group for the lesson *Potential* and *Implementation* means listed in Table 4.11. Implementation means from the lesson observations exhibited an overall increase of 0.46 between Fall and Spring. However, results of the Mann-Whitney tests indicate that this increase is not statistically significant (z = 1.21; p = .11 [one-tailed]); increases in *Implementation* scores for lesson observations were not significant between any of the data collections.

No significant differences were found between the contrast group and the ESP teachers' Fall scores for *Potential* (z = 0.81; p = 0.21 [one-tailed]) or for *Implementation* (z = 0.67; p = .25

[one-tailed]). This indicates that prior to participation in the ESP workshop, ESP teachers used similar levels of tasks and implemented them in similar ways as teachers in the contrast group. Conversely, ESP teachers' Spring lesson observation scores were significantly higher than the contrast group in *Potential* ($z = 2.15$; $p = 0.02$ [one-tailed]) and in *Implementation* ($z = 1.87$; $p = 0.03$ [one-tailed]). Following their participation in the ESP workshop, ESP teachers were selecting and implementing high-level tasks more frequently than their counterparts in the contrast group.

Table 4.12.
Comparison of Lesson Observation Implementation Scores

	Differences in Means	
	Potential	*Implementation*
Comparisons between Data Collections		
Fall vs. Winter	0.55*	0.41
Winter vs. Spring	0.05	0.05
Fall vs. Spring	0.50*	0.46
Comparisons to Contrast Group		
Fall vs. Contrast	0.28	0.26
Spring vs. Contrast	0.78*	0.71*

*Significant at $p < 0.05$.

Due to sample size, observational comparisons were used to identify changes in the number of lesson observations that began with high-level tasks (i.e., a score of 3 or 4 for *Potential*) and remained high-level throughout the lesson (i.e., a score of 3 or 4 for *Implementation*). Data in Table 4.13 reports the number of paired *Potential*-vs.-*Implementation* scores for lesson observations in each data collection, categorized as High-High (H-H) (i.e., lesson observations with a score of 3 or 4 for *Potential* and for *Implementation*), High-Low (i.e., lesson observations with a score of 3 or 4 for *Potential* and a score of 1 or 2 for *Implementation*), or Low-Low (i.e., lesson observations with a score of 1 or 2 for *Potential* and for *Implementation*). No lesson observations in any data collection increased from a low-level score for *Potential* to a high-level score for *Implementation*. Comparisons between data collections indicate that three teachers changed scores from L-L to H-H between Fall and Winter. No change in quantities occurred between Winter and Spring.

Changes in the number of lesson tasks implemented at a high-level between the Fall and Winter observations are reflected in factors identified on the IQA Lesson Checklist, which indicated an increase in 1) teachers holding students accountable for high-level products and processes; 2) teachers providing consistent press for explanation and meaning, and 3) teachers providing encouragement for students to make conceptual connections. Winter Lesson Checklists also indicate a corresponding decrease in the occurrence of 1) teachers providing a set procedure for solving the task; 2) the focus shifting to procedural aspects of the task or on correctness of the answer rather than on meaning and understanding; 3) feedback, modeling, or examples being too directive or not leaving any complex thinking for the student; and 4) students not being pressed or held accountable for high-level products and processes or for explanations and meaning. The changes identified between Fall and Winter were sustained between Winter

and Spring, though no new changes or patterns emerged. The contrast group more closely reflected the patterns and factors of implementation exhibited by the ESP teachers in the Fall. Hence, following their participation in ESP, the ESP teachers were more likely to exhibit classroom factors that maintained high-level cognitive demands and less likely to exhibit classroom factors that reduce high-level cognitive demands (Stein, et al., 1996) than a) prior to their participation in the ESP workshops and b) than a group of teachers who did not participate in the ESP workshops.

Table 4.13.
Comparison of High-level vs. Low-level Potential and Implementation for Lesson Observations

	Potential- Implementation		
	High-High	High-Low	Low-Low
Project Group (n = 11)			
Fall	4	2	5
Winter	7	2	2
Spring	7	2	2
Contrast Group (n = 10)	2	2	6

4.4. Role of the ESP Professional Development Workshop

This section addresses research Question #4:

> What changes (or lack thereof) in teachers' knowledge, selection, or implementation of cognitively challenging tasks can be reasonably associated with individual or group experiences in the professional development sessions?

From Fall to Spring, ESP teachers significantly improved their ability to characterize high and low-level tasks, increased their use of high-level tasks as the main instructional tasks in their own classrooms, and improved their ability to maintain the cognitive demands of a high-level tasks as evident in student' work. ESP teachers were also significantly more knowledgeable of the cognitive demands of mathematical tasks, more likely to use high-level tasks as the main instructional tasks in their own classrooms, and better able to maintain high-level cognitive demands during instruction than a contrast group of secondary mathematics teachers who did not participate in the ESP professional development workshop. This section will identify the opportunities within the ESP workshop that may have generated the observed changes in ESP teachers' knowledge, selection, and implementation of cognitively challenging tasks and provided opportunities for ESP teachers to consider the use and implementation of mathematical tasks in their own classrooms.

Figure 4.1 portrays all of activities conducted within the six ESP workshops, highlighted to identify the activities that explicitly addressed 1) the level of cognitive demand of mathematical tasks (yellow), 2) the selection and/or implementation of high-level mathematical tasks (blue) within the context of practice-based professional development materials, and 3) the selection and/or implementation of high-level mathematical tasks in teachers' own classrooms

(orange). Teachers' opportunities to consider the selection and implementation of high-level tasks were coded as a single category since ideas about task selection were either implicit or intertwined in discussions about task implementation in ways that made task selection impossible to code as a separate category. In contrast, the activities and discussions pertaining to the use of high-level mathematical tasks in teachers' own classrooms were very distinct from the activities and discussions pertaining to the professional development materials, and preserving this distinction appeared valuable. Table 4.14 lists the times spent on the three categories of activities, and the paragraphs that follow describe each category in greater detail.

4.4.1. Discussions about the Level of Cognitive Demand of Mathematical Tasks.

ESP teachers often participated in discussions that engaged them in considering the cognitive demands of mathematical tasks. Such discussions were most prominent in the first two sessions, encompassing approximately 90 minutes in Session 1 and 42 minutes in Session 2. Time spent explicitly discussing the cognitive demands of mathematical tasks decreased substantially in Sessions 3 through 6.

The discussion of cognitive demands initiated in Session 1 consisted of a comparison between two tasks similar in mathematical content but very different in cognitive demand (i.e., the Martha's Carpeting Task and the Fencing Task [Stein, et. al., 2000]). Following this comparison, teachers discussed their individual criteria for categorizing tasks as "High-Level" or "Low-Level" on the pre-workshop task sort, and collectively constructed a set of criteria for categorizing task demands as high-level or low-level. In Session 2, teachers were introduced to the Task Analysis Guide (TAG) (see Figure 2.1), and their analysis of the

Session 1: Oct. 2, 2004	Session 2: Nov. 6, 2004	Session 3: Jan. 8, 2005	Session 4: Feb. 5, 2005	Session 5: Mar. 5, 2005	Session 6: May 7, 2005
Introductions & Data Collection	Introducing Levels of Cognitive Demand and The Mathematical Tasks Framework	Reflecting on Sessions 1 & 2	Why Cases?	Case Stories II: Storytelling through Student Work. What did students' work tell about maintaining high-level cognitive demands during the lesson?	Case Stories III: How did assessing & advancing questions influence the enactment of the task?
Solving "Martha's Carpeting" & the "Fencing" Tasks	Solving the "Linking Fractions, Decimals, & Percents" Task	Multiplying Monomials and Binomials: Developing the area model of multiplication	Case Stories I: Reflecting on Our Own Practice. How did the factors of scaffolding and press play out in the lesson?		Planning the "Sharing and Discussing" Phase of a Lesson: Selecting and ordering presentations
Comparing Martha's Carpeting Task & the Fencing Task: How are they same and/or different?	Reading & Discussing the Case of Ron Castleman: Similarities and differences between 2nd and 6th period. Do the differences matter?	Solving the "Multiplying Monomials & Binomials" Task with Algebra Tiles	Solving the "Extend Pattern of Tiles" Task	Focusing on the "Exploring the Task" Phase of a Lesson: What questions would you ask to assess and to advance students' understanding?	Introducing the "Thinking Through a Lesson" Protocol
Categorizing Mathematical Tasks: The Task Sort	The Factors and Patterns of Maintenance & Decline	Reading & Discussing the Case of Monique Butler: What did MB *want* her students to learn and what *did* they learn?	Analyzing Student Work on the Extend Pattern of Tiles Task: Which show greatest/least understanding?	Planning a Whole-Group Discussion: What responses would you share & why?	Solving "Double the Carpet" Task
Data Collection, Paperwork	Data Collection, Paperwork	Connecting to Own Teaching: Discuss factors that influenced your lesson			Data Collection, Paperwork
Identify a task from your data collection that you would like to change/adapt/improve in some way.	*Plan, Teach and Reflect on a lesson involving a high-level task: identify factors at play in your lesson and factors you want to work on this year*	*Plan, teach and reflect a lesson using a high-level task. In what ways did you make progress on the factor you have chosen? What do you still need to work on?*	*Plan, Teach and Reflect on a lesson involving a high-level task: before and after, complete the chart on factors and expectations. Bring in student work.*	*Plan, Teach and Reflect on a lesson involving a high-level task. List questions to assess & advance Ss learning. Bring in list of questions and student work.*	*Plan, Teach and Reflect on a lesson involving a high-level task. Use the TTAL to plan and reflect on the "Sharing & Discussing" phase of your lesson.*
Color Code: Opportunity to learn about 1) level of cognitive demand of mathematical tasks Opportunity to learn about selection or implementation of high-level tasks Opportunities to consider use of high-level tasks in own classroom					

Figure 4.1 . ESP professional development activities for Cohort 2 (2004-2005).

Table 4.14.

Opportunities in the ESP Sessions for Teachers to Consider the Level and Use of Mathematical Tasks

	Level of Cognitive Demand of Mathematical Tasks		Selection and/or Implementation of High-Level Tasks		Use of High-Level Tasks In Own Classroom	
Session	Activity	Time	Activity	Time	Activity	Time
1	Comparing Martha's Carpet vs. Fencing Task	31:45	How did the facilitator support your learning (Fencing Task)?	2:45	Assignment 1	NA
	Categorizing Mathematical Tasks	57:20				
2	Introduction of LCD	4:20	How did the facilitator support your learning (FDP Task?)	23:20	Assignment 2	NA
	Analyze LCD of "Linking Fractions, Decimals, & Percents" (FDP) Task	37:35	Implementation of FDP task in "The Case of Ron Castleman"	1:15:30		
			Introduction of MTF	3:20		
			Presentation of Factors, Patterns	28:30		
3	LCD of Alg. Tiles task	2:15	Advantages of using higher-level task (Alg. Tiles)	8:30	Discussion of what they did/thought differently	10:30
			Implementation of Alg. Tiles task in "The Case of Monique Butler"	1:06:25	Discussion of tasks that they used/shared	9:30
					Share implementation of own high-level task	44:35
					Assignment 3	NA

Table 4.14 (continued).

	Level of Cognitive Demand of Mathematical Tasks		Selection and/or Implementation of High-Level Tasks		Use of High-Level Tasks In Own Classroom	
Session	Activity	Time	Activity	Time	Activity	Time
4	LCD of Extend Pattern of Tiles Task	18:15	Analyzing Student Work on EPT Task	1:08:20	Case Stories 1	1:18:05
					Assignment 4	NA
5			Assessing and Advancing Students' Understanding	1:12:45	Case Stories 2	1:27:30
					Assignment 5	NA
			Planning a Whole-Group Discussion	1:15:05		
6	LCD of Double the Carpet Task	3:50	Planning a Whole-Group Discussion (continued)	42:30	Case Stories 3	1:04:55
			Intro to the TTAL	23:45	Discussion of how assessing and advancing questions influenced the enactment of the task	19:45
			Planning a Lesson around the Double the Carpet Task	36:10		
					Assignment 6	NA

cognitive demands of mathematical tasks was enhanced from a dichotomous categorization of high-level vs. low-level to a more fine-grained distinction between specific types of high-level tasks (i.e., doing mathematics and procedures with connections) and specific types of low-level tasks (i.e., procedures without connections and memorization). Teachers used the TAG to categorize mathematical tasks throughout the remainder of the ESP sessions.

How do teachers' opportunities to learn about the cognitive demands of tasks within the ESP workshop compare to the nature of the changes in teachers' knowledge of the cognitive demands of mathematical tasks? Recall that a specific set of criteria for high- and low-level tasks appeared in teachers' post-workshop task sort responses that were not present in teachers' pre-workshop task sort responses. Specifically, the criteria noted on the post-test were: *specific level of cognitive demand*, *presence of a stated or implicit procedure*, opportunities for *connections*, use of *multiple representations*, opportunities for *generalizations*, and opportunities for *multiple solution methods*. These criteria were prominent in discussions of the cognitive demands of mathematical tasks within the ESP workshop. For example, teachers consistently identified multiple solution strategies (or "open-ended") as a feature of high-level tasks and consistently associated the presence of a prescribed procedure as a feature of low-level tasks. Discussions of whether tasks were at the level of "doing mathematics" or "procedures with connections" focused on whether a procedure was suggested by the task or whether the task allowed for multiple strategies. If a task prescribed a procedure, teachers then addressed whether the procedure provided students with opportunities to make mathematical connections or whether students were applying a rote procedure with no connection to meaning. Two other features of tasks arose more than once during discussions, the prompt for an *explanation* and the *use of diagrams*, but these features did not emerge as prominent criteria on ESP teachers' post-

workshop task sort responses. Figure 4.2 identifies when the features noted in this paragraph (in italics) arose during the ESP sessions, and the features representative of changes in ESP teachers' task sort responses are denoted in Figure 4.2 in bold.

The criteria that emerged on the post-test were often explicitly modeled by the facilitators during discussions of teachers' own work on mathematical tasks (i.e., "Were you surprised by all of the different strategies?" [video transcript, Session 2, 11/06/04]; "What is different about Iris and Randy's strategy?" [video transcript, Session 4, 2/05/05]; "How does the equation connect to the diagram?" [video transcript, Session 6, 5/07/05]). In each session, the time spent solving mathematical tasks was not coded as providing explicit opportunities for teachers to consider the cognitive demands of mathematical tasks. Only the portions of the task discussions that explicitly addressed the level, selection or implementation of high-level tasks were earmarked for Table 4.14. Arguably however, engaging with a high-level task may have allowed teachers to implicitly attend to features and characteristics of the task that provide opportunities for high-level thinking and reasoning. In five of the six sessions, teachers were asked to consider and discuss the cognitive demands of the tasks they had engaged in solving (see Tables 4.14 and Figure 4.2). In each of these discussions, teachers explicitly identified high-level features of their own work on the task as characteristics that gave the task high-level cognitive demands. Consider the whole-group discussion of the Fencing Task in Session 1. Teachers presented and discussed multiple solution strategies and multiple representations, and the facilitator made explicit moves to foster connections between strategies (i.e., "Do you see any connections between Randy and Dave's solutions?" [video transcript, Session 1, 10/02/04]) and between representations (i.e., "What is it about the table that gives you a clue about the graph?" [video transcript, Session 1, 10/02/04]). During the comparison of Martha's Carpeting Task and the Fencing Task, teachers identified

	Specific use of TAG	**Multiple solution strategies**	**Multiple representations**	**Generalizations**	**Connections[abc]**	**Prescribed or implicit solution method**	Explanation	Use of diagrams
Session 1								
Teachers' work on the Fencing Task		X	X		a, b			
Martha's Carpet vs. Fencing Task		X	X	X	a, b, c			
Categorizing Mathematical Tasks		X		X	c	X	X	
Session 2								
Introduction of TAG	X							
Teachers' work on "Linking…" Task		X			c			X
Analyze LCD of "Linking…" Task	X	X			c	X	X	X

Note: Headings in bold indicate changes noted in teachers' post-workshop task sort criteria
[a]a indicates connections between strategies
[b]b indicates connections between representations
[c]c indicates connections between mathematical concepts

Figure 4.2. Features of the level of cognitive demand of tasks that arose during ESP discussions.

Figure 4.2 (continued)

	Specific use of TAG	**Multiple solution strategies**	**Multiple representations**	**Generalizations**	**Connections[abc]**	**Prescribed or implicit solution method**	Explanation	Use of diagrams
Session 3								
Teachers' work on Alg. Tiles Task			X		b, c			X
LCD of Alg. Tiles task	X							
Session 4								
Teacher's work on EPT Task		X	X	X	a, b, c			X
LCD of EPT Task	X	X			c	X		X
Session 6								
Teachers' work on "Double …" Task		X	X	X	a, b, c			X
LCD of Double the Carpet Task	X						X	

Note: Headings in bold indicate changes noted in teachers' post-workshop task sort criteria
[a]a indicates connections between strategies
[b]b indicates connections between representations
[c]c indicates connections between mathematical concepts

multiple strategies, multiple representations, and connections between strategies and representations (i.e., features were prominent in teachers' own work on the task) as characteristics that made the Fencing Task "different" from Martha's Carpeting. Comments from two participants during the comparison of the tasks illustrates that teachers were drawing on their experiences in solving the tasks as learners::

> Michelle: I actually learned something with doing (the Fencing) task. We all solved Martha's Carpeting the same way. But the Fencing task, the discussion that was going on at our table, we started getting into the graphs and the parabola, and through somebody else's solution at my table that I didn't think of myself, I actually started making those connections.
>
> Nellie: I agree with learning something. I liked seeing all the different ways, especially the Algebra 2 and calculus. It really made me make connections. (video transcript, Session 1, 10/02/04).

Similarly, while solving the Linking Fractions, Decimals, and Percents Task (Stein, et. al., 2000) in Session 2, teachers were provided with resources to enable them to create a variety of strategies and were prompted to use the diagram to explain their thinking. In the discussion of the level of cognitive demand of the task, opportunities for multiple strategies and the requirement to make connections to the diagram were noted as characteristics that made the task high-level. The discussion of teachers' work on the EPT task in Session 4 was characterized by using diagrams and making generalizations, which emerged as prominent ideas in the discussion of the level of cognitive demand of the EPT task. Teachers' engagement with tasks as learners appears to have influenced their knowledge of the level of cognitive demands of mathematical tasks.

Hence, at several points throughout the ESP workshop, teachers had opportunities to increase their knowledge of the cognitive demands of mathematical tasks. Furthermore, the criteria for categorizing high- and low-level tasks that emerged on ESP teachers' post-workshop task sort responses were frequently addressed during discussions of the level of cognitive demands of tasks and during teachers' own work on mathematical tasks.

4.4.2. Discussions about the Selection and Implementation of High-Level Tasks

As identified in Table 4.14, discussions pertaining to the selection and implementation of high-level mathematical tasks were central features of Sessions 2 through 6. In Session 2, ideas about task implementation were initiated with a discussion of how the facilitator supported the teachers' own learning during their engagement with the Linking Fractions, Decimals, and Percents task. Teachers then considered the implementation of the same task in The Case of Ron Castleman (Stein, et al., 2000) In this discussion, teachers compared and contrasted two lessons in Ron Castleman's classroom – one lesson in which the high-level demands of the Linking Fractions, Decimals, and Percents task declined during implementation and another lesson in which the high-level task demands were maintained. The comparison allowed teachers to distinguish between features of instruction that sustain vs. diminish students' opportunities to engage with high-level task demands throughout an instructional episode. Following the discussion of the case, teachers were introduced to: 1) the Mathematical Tasks Framework (MTF) (see Figure 2.2); 2) a set of classroom-based factors that influence the maintenance and decline of high-level cognitive demands (see Figure 2.3); and 3) patterns of maintenance and decline of high-level tasks. These factors and patterns were used in Session 3 to analyze the teaching and learning that occurred in The Case of Monique Butler (Stein, et al., 2000), and in

subsequent ESP sessions to assess the implementation of high-level tasks within practice-based professional development materials and within teachers' own classrooms. In Sessions 4 through 6, the implementation phase of the MTF was dissected into the components of "Supporting Students' Exploration of the Task" and "Sharing and Discussing the Task." Teachers considered each component in detail; assessing and advancing students' mathematical understandings through the use of questioning (Session 4 and 5), and orchestrating whole-group discussions characterized by the presentation of a variety of strategies, the ordering of those strategies so as to surface the mathematical ideas and foster connections, and questioning by the teacher to advance students toward the target mathematical goals (Sessions 5 and 6). In a capstone activity at the close of Session 6, teachers collaboratively planned a lesson using a lesson planning tool closely aligned with ideas about implementing high-level tasks in ways that preserve high-level task demands and foster thinking and reasoning amongst students (i.e., the Thinking Through a Lesson Protocol [Hughes & Smith, 2004]). Overall, the set of discussions pertaining to the selection and implementation of high-level tasks began by viewing task implementation from the perspective of the learner (i.e., considering how the facilitator supported their learning). This perspective was enhanced when viewed through the lens of the MTF and the factors that support vs. diminish high-level implementation. From this viewpoint, specific components of implementation were then magnified and scrutinized in greater detail.

How do teachers' opportunities to engage with ideas about the selection and implementation of high-level tasks within the ESP workshop compare to the changes in their selection and implementation of high-level task in their own classroom? Teachers exhibited a significant increase in the use of tasks with higher level cognitive demands, which they explicitly attributed to their experiences in the ESP workshop:

Michelle: "…since we've been doing all this focusing on tasks…, I've started taking a closer look at tasks that I give my students to do, and how I could make them more of a high-level. …" (video transcript, Session 3, 1/08/05);

Dan: "I used to use this (task) as bonus. Many of the problems that I used to use as bonus, since this program, I'm using them as a whole lesson" (video transcript, Session 6).

At the beginning of Session 3, all 18 ESP teachers indicated in their journals that they were using or considering the use of more high-level tasks in their classrooms. Seven teachers stated that *they were thinking about using* more high-level tasks (i.e., "I realized that I rarely use high-level tasks in my (middle school math) class. I thought that I should start incorporating them on a more regular basis" [Natalie, journal entry, Session 3, 1/08/05]) and eleven teachers indicated that *they had begun to use* more high-level tasks and/or were thinking about the implementation of the high-level tasks already used in their classroom (i.e., "focus on the kinds of questions I ask…listen to students' responses to see if their ideas could lead to further discussions" [Cathy, journal entry, Session 3, 1/08/05]). During the discussion, two teachers commented that, since their participation in ESP, they had begun using different types of tasks as warm-up tasks (i.e., tasks that elicited multiple strategies), and three teaches indicated that they had used the tasks from Sessions 1 and 2. Hence, ESP teachers appear to be using or considering the use of a greater number of high-level tasks following Session 2. At this point in the workshop, they had engaged with high-level tasks as learners, they had categorized and developed criteria for high-level and low-level tasks, and they had been exposed to research on the influence of high-level task on students' learning (i.e., Stein & Lane, 1996). The Winter data collection occurred in the month following Session 3, and the gains in task means between Fall and Winter reflect teachers' self-reported use and intended use of higher-level tasks. By Spring, task means and the number

of high-level tasks per teacher both increased significantly, providing evidence that teachers (16 of the 18) began using or continued to use high-level tasks for instruction in their own classrooms throughout their participation in ESP.

Teachers also exhibited an increase in their ability to maintain high-level cognitive demands in students' work, as student-work implementation scores increased significantly from Fall to Spring. Teachers' support of students' work was often addressed within the ESP sessions. Teachers engaged in reading, analyzing, and discussing two narrative cases in which the maintenance of high-level task demands during implementation was examined very closely. In each case, the ways in which the teachers supports and scaffolds students' work on the task greatly influences students' opportunities to engage with high-level thinking and reasoning, or conversely, to engage with procedures without connection to mathematical meaning and understanding. ESP teachers were asked to extract general lessons learned from these cases that applied to teaching and learning mathematics more broadly than the specific task and lesson featured in the case. The general lessons learned may have been applied to task implementation in their own classroom, and in doing so, generated the observed increase in student work implementation means from Fall to Spring.

The use of questioning emerged in Sessions 1, 2, 4, and 6 as participants noted instructional moves modeled by the ESP facilitators that supported teachers' own engagement with high-level mathematical tasks. The use of questioning was also an explicit focus of the discussions in Sessions 5 and 6, as teachers considered student work samples and constructed questions that assessed and advanced the student's mathematical understandings. Features of each type of question (i.e., assess and advance) were then generalized to apply beyond the specific task and student work under examination. In the evaluations for Session 5, eleven

teachers commented on the use of assess and advance questions (i.e., "I am planning to assess my students by questioning to find level of understanding so I know what advancing questions to ask" [written artifacts, Session 5, 3/05/05]). Improvements in teachers' use of questioning, as a tool for supporting students work, may account for the increase in student work implementation scores.

As stated in the previous section, of the time spent solving mathematical tasks, only the portions of the discussions that explicitly addressed the level, selection or implementation of high-level tasks were coded for Table 4.14. As argued earlier, teachers' experiences in solving high-level tasks and in discussions on the cognitive demands of tasks may have provided implicit opportunities for teachers to consider the selection and implementation of high-level tasks in their own classrooms. Facilitators modeled the pedagogy of good instruction (Smith, 2001), modeled instructional factors that maintain high-level cognitive demands, and made instructional moves that supported the development of the mathematical ideas. Comments and journal entries provide evidence that teachers were attending to implementation as they engaged in solving tasks and discussing the cognitive demands of tasks:

Dan: "Interesting to watch how (the facilitator) responded to all of the different answers. You need to be prepared for anything" (journal entry, Session 1, 10/02/04).

Kathy: "I saw a good model of questioning. (Facilitator) had a technique of asking one group for an answer and asking another group to explain the answer. This got more people involved" (written artifacts, Session 1, 10/02/04).

Kelsey: "..it might depend on how the teacher implements it in class. *Certainly, the way we did it here*, we used representations, made connections between those representations… But

at the same time, if we only looked at one solution, and didn't take the time to discuss some more, then it could go either way" ([emphasis added] video transcript, Session 2, 11/06/04).

To make this modeling explicit, teachers were asked to reflect on how the facilitator supported their learning. Figure 4.3 provides the list of teachers' responses when this question was posed in Session 2.

Teachers engaged in many activities which enabled them to consider the selection and implementation of high-level tasks within the context of tasks, student work, and classroom episodes featured in the practice-based professional development materials used in the ESP sessions. Teachers' opportunities to consider the use of high-level tasks in their own classroom will be discussed in the following section.

4.4.3. Discussions about Teachers' Use of High-Level Tasks in their Own Classrooms

Sessions 3 through 6 provided teachers with opportunities to discuss and receive feedback on the use of high-level tasks in their own classrooms. In the assignment prior to each session, teachers were asked to select and implement a high-level task during a lesson in their own classroom, and to reflect on certain aspects of their implementation. The assignments are identified in italics in the bottom row of Figure 4.1. Note that teachers were asked to plan and teach a lesson based on a high-level task following Sessions 3 and 5, which coincided with the Winter and Spring data collections, respectively. Teachers' data collections were compared to the written artifacts submitted for their assignments to determine whether the data collection included the tasks and/or student work from the lesson that teachers had used for their assignments. In the Winter, five of the 16 teachers (31%) who submitted data collections included the task used in their Session 3 assignment as one of the 5 main instructional task in

their data collection. In the Spring, four of 14 teachers (29%) submitted the task and/or student work from their Session 5 assignment. This results in 5 of 80 tasks (6%) in the Winter data collection and 4 of 70 tasks (6%) in the Spring data collection stemming from teachers' assignments for the ESP workshop. Hence, the assignments only minimally intersected with teachers' data collections. Interestingly, the tasks used in the ESP workshop also did not appear in ESP teachers' data collections. Only one ESP teacher submitted tasks and student work in the data collection that could be directly connected to tasks used in the ESP workshop; following the "Multiplying Monomials and Binomials" tasks using algebra tiles in Session 3, the teacher constructed a week-long exploration in which students used algebra tiles to multiply and factor algebraic expressions.

As was characteristic of other types of activities in the ESP sessions, the use of high-level tasks and reflections upon their implementation became increasingly fine-grained as the sessions progressed. The assignment following Session 2 asked teachers to select and implement a high-level task, and to reflect on which factors influenced their implementation of the task. Teachers returned to Session 3 with their reflections, discussed the lesson with their peers, and co-constructed ways they might work on the factors in their classrooms. Session 4 initiated the activity of teachers telling "case stories" (Ackerman, Maslin-Ostrowski, & Christensen, 1996) about their lessons, which continued in sessions 5 and 6. Case stories are a structured format for teachers to share their teaching practice with a small group of colleagues and for colleagues to provide feedback. Teachers were asked to teach a lesson using a pedagogical 'tool' highlighted within the previous ESP session and were provided with specific prompts to reflect on the teaching and learning that occurred in the lesson.

What supported your learning?

- Sought out a broad range of solution methods (non-routine) (visited groups)
- Individual time
- Careful selection of methods – ordering & connections – flowed
- Had an idea of where she wanted to go
- Time sufficient
- Resources – markers, grids
- No calculators provided
- Posed questions to help group look at things different
- Great question <u>why</u> during discussion [fractions vs. ratios]
- Make connections/use evidence
- Having solutions presented
- Make sense of others' solutions
- Visual display
- Created comfortable atmosphere

Figure 4.3. **Chart of responses to "What did the facilitator do to support your learning?" (Session 2; 11/05/04).**

Teachers returned to the next ESP session with their written reflections and any evidence or artifacts (i.e., student work, lesson plans, transcribed or paraphrased interactions, video- or audio-taped segments of the lesson) to tell the 'story' of their lesson. In Session 4, teachers' stories were based on their own reflections. In Session 5, teachers supplied student work to enhance the story and provide evidence of the implementation of the task. For Session 6, student work and lists of teachers' questions and students' responses served as evidence of the lesson.

The assignments and the case stories that accompanied them provided opportunities for ESP teachers to use the ideas and tools presented in the previous ESP session in their own classroom, and to discuss their use of the ideas and tools with colleagues and with members of the ESP team.

Discussions of selection and implementation of mathematical tasks also appear to have provided teachers with opportunities to consider the use of high-level tasks in their own classrooms. Teachers spontaneously reflected from the case discussion to their own classroom and teaching practices (Walen & Williams, 2000) (i.e., "Makes me think about the type of questions that I pose" [Maureen's journal entry, Session 2, 11/06/04]). In addition to their spontaneous reflections, teachers were frequently asked to make explicit connections to their own practice (i.e., how to create opportunities for high-level thinking for their students, general lessons learned from the narrative cases in Sessions 2 and 3).

Experiences in applying the ideas and tools of ESP, support for their continued improvement in implementing high-level tasks, and reflections from the professional development to their own classroom may have influenced teachers' increased selection and implementation of high-level tasks in their own classrooms as evidenced by the data. The following section will examine the ways in which specific teachers exhibited change, and how these specific teachers are characteristic of other teachers with similar patterns of growth.

4.4.4. Portraits of Instructional Change for Selected Teachers

From the qualitative analysis, three teachers were selected to portray the nature of instruction in their classrooms at different points in time with respect to the selection and implementation of cognitively challenging tasks. This section will provide accounts of three ESP teachers' participation in the ESP sessions and the ways in which they exhibited (or did not

exhibit) instructional change. The case studies that follow feature Randy, who improved his knowledge, selection and implementation of high-level tasks; Nellie, who was using high-level tasks prior to her participation in ESP but improved in knowledge and implementation of high-level tasks; and Cara, who exhibited growth only in her knowledge of cognitive demands. These teachers were chosen because they are representative of others teachers with similar patterns of change and because they are similar to each other in age, preservice MAT program, years of teaching experience, and school demographics. In fact, several similarities exist between Nellie and Cara that makes the differences in their participation in ESP and in their patterns of change very surprising. This section concludes by summarizing the characteristics of Randy, Nellie, and Cara that are representative of other teachers with similar patterns of change.

4.4.4.1. A closer look at a teacher who improved in knowledge, selection and implementation of high-level tasks

During his participation in ESP, Randy was teaching in a large affluent suburban school and had 7 years of teaching experience. He joined ESP along with three of his colleagues on the recommendation of the mathematics coordinator at his school. Similar to Randy, his colleagues had between 3 and 9 years of teaching experience. Randy and his colleagues were considered to be a thoughtful and dedicated group of young teachers who would consider the ESP professional development workshop to be a valuable learning experience. Randy submitted data for a 10th-grade mathematics class that integrated algebra and geometry and used a reform-oriented textbook series. He was simultaneously piloting units from the algebra and geometry texts in another reform-oriented series that his school was considering adopting in the following school year.

Randy's score of 21 on the pre-workshop task sort was below the group average of 24.21, and was the 5th lowest of the 19 ESP teachers. Hence, Randy can be considered to have entered the ESP workshop with low knowledge of the cognitive demands of mathematical tasks. Randy's Fall task mean was 2.2, with one high-level task present in the collection of five main instructional tasks. Randy's task mean was below the Fall average of 2.54, and he used fewer high-level tasks than the average of 2.2 per teacher. Why were Randy's task scores low even though he was using a reform-oriented textbook? On 4 of the 5 days in the Fall task collection, the main instructional tasks used in Randy's classroom were obtained from a supplemental resource of practice worksheets that accompanied the text. These worksheets were intended for use as a skills review, and were coded at the level of procedures without connections (i.e., a score of 2). The textbook itself, which contains opportunities for explorations and sense-making, was not utilized as a source of main instructional tasks in Randy's classroom.

Two tasks in Randy's Fall student work collection were rated as high-level in *Potential* (i.e., a score of 3). Randy obtained both of these tasks from outside resources, and neither task was used as a main instructional task; one was a brief warm-up activity (though it was rich enough to have taken an entire class period) and the other was homework. As implemented in students' work, both tasks declined to procedures without connections (i.e., a score of 2) for *Implementation*. Students' work for solving both of the high-level tasks exhibited one strategy for solving the problems, the use of a systematic guess-and check procedure followed by substituting numbers into formulas. Only the procedural steps appear in students' written work, perhaps because students did not have enough time to engage with high-level task demands for the task that was used as a warm-up and did not have press for explanation and meaning for the task that was completed as a homework assignment.

Similarly, Randy's Fall observation was rated as procedures without connections, as both the *Potential* and *Implementation* of the task received a score of 2. The lesson consisted of a review of factoring that consumed approximately 30 of the 40 minutes of class time. Randy modeled how to factor a quadratic equation, and then presented students with similar problems to practice with a partner. As students worked, Randy selected students to write only their solutions on the SmartBoard. He then conducted a whole-group discussion in which he elicited from students the steps for factoring several of the equations. Randy closed the lesson by telling students why they had reviewed factoring and how it would be important in the upcoming lesson. In the post-lesson interview, Randy noted that he had planned to spend far less time on factoring (it was intended as a 10-minute warm-up activity) and to incorporate a problem-solving task as the main instructional task. Due to time constraints, the problem-solving task was given to students as homework.

Winter data presents a different portrait of instruction and learning in Randy's classroom. Randy's task mean was 3.1, with 4 of the main instructional tasks scored as high-level. This exceeds the ESP teachers' Winter average of 3.38 high-level task per teacher and Winter task mean of 2.93. What was different about Randy's use of tasks in the Winter data collection? On two days, he incorporated activities from the textbook that involved high-level thinking and reasoning. On two other days, he incorporated high-level tasks that he and a colleague had created. In contrast to Fall, he did not center instruction on the skill-based practice worksheets that accompanied his text. Instead, high-level tasks from the text and from outside resources were used as the main instructional tasks and consumed the majority of instructional time in his classroom (25-30 minutes per lesson).

In the Winter data collection, implementation was beginning to improve, as well. Two of the three tasks in the student work collection were rated as high-level for *Potential.* Both tasks were at the "procedures with connections" level of cognitive demand. For one task, the potential connections were not evident in students' work (students' work consisted of formulas evaluated at different numeric values), and the *Implementation* declined into procedures without connections (i.e., a score of 2). In contrast, the other high-level task was maintained as procedures with connections, as students were able to write general rules for simplifying exponents with like bases and connections between the rules and the meaning of exponents were implicitly evident in students' written work (i.e., students used the expanded form of exponents to derive the rules, but did not provide explicit explanations of how the rules were derived, why the rules made sense, or why the rules would always work). During the Winter lesson observation, Randy used <u>and</u> was able to maintain a high-level task that required explicit connections to meaning (i.e., a procedures with connections task at the score level of 4 for both *Potential* and *Implementation*). During the lesson, students were asked to generalize the rules for simplify exponents with like bases after examining several numeric examples. Students were resistant to expanding the exponents into their factors, and wanted to jump ahead and simply apply the rules. Randy consistently held students accountable for expanding the exponents, forming general rules of their own before looking ahead, and examining why the rules worked. As he visited the groups, he reminded students, "<u>Yes</u>, you need to write it out," and "don't move ahead and just write (*the answer*)." In response to students' complaints that the problems were "tedious" and "a waste of a step," Randy responded, "Perhaps there is a point to it," and "Show us that this property works. I think they want you to see something." As students discussed their findings on the first set of examples, Randy asked, "Okay, so what was the point of writing it

out?" Students were able to respond that, for the example $2^4 * 2^3$, you "write seven 2s," and that "when you are multiplying with exponents just add the exponents together," and following press by Randy, another student added, "when the base is the same." Small and large group discussions were characterized by questions from Randy that held students accountable for writing the exponents in expanded form and pressing students to make sense of the rules for simplifying exponents with like bases. Students were resistant to engage in high-level thinking; they pressed the teacher to proceduralize the task and allow them to apply the rules with no connection to meaning. Based on their Fall student work collection and observation, this is the way typical work in math class was conducted. By the Winter data collection, Randy was attempting to implement new expectations, and students were not yet accustomed to working this way.

Randy continued to improve his implementation of high-level tasks in the Spring data collection. All three student work tasks began as high-level and were implemented at a high level. Similarly, the Spring observation used a procedures with connections task at a score level of 4 as the main instructional task, and Randy maintained the task at a level 4 during instruction. In this lesson, students described the graphs of given functions by comparing them to one of nine parent functions the class had studied over the course of the year. Only the equations of the parent functions were listed on the chalkboard. Students worked with partners as Randy circulated. During the whole group discussion, Randy asked students to identify the type of function represented by the parent function they had chosen. When students provided an answer, he continued to press for meaning (i.e., How did you know?; What gave it away?; Why is it [*that function*]? [Spring observation notes, 5/12/2005]). He also asked questions that enabled students to make connections between the graphs of a given family of functions (i.e., Where would $y = 3^x$

or $y = 5^x$ be relative to 2^x and 10^x?) and to make sense of the domain and range of a given family of functions (i.e., Why does $y = \sqrt{x}$ stop at (0, 0)?; Why can the next graph ($y = \sqrt[3]{x}$) extend past 0?).

In summary, Randy improved his use of high-level tasks as the main instructional tasks in his classroom between the Fall and Winter data collections. In the Winter data collection, Randy was beginning to implement tasks at a high-level, and these improvements continue into the Spring. In the overall data from ESP teachers, Randy was 1 of 4 teachers (of the 11 who were observed) that maintained a task at level 4. In fact, there were only 7 instances of *Implementation* scores of 4 within the 33 observations, and Randy contributed 2 of them.

In what ways might the ESP workshop have influenced the changes in Randy's practice? To begin with, Randy increased his knowledge of cognitive demands of mathematical tasks between Fall and Spring. This is evidenced by the 8-point gain in Randy's post-workshop task sort score, which increased from 21 to 29 and placed Randy slightly above the mean of 28.79 and 7th highest of the 19 ESP teachers. Increased knowledge of the cognitive demands of mathematical tasks may have influenced the selection and implementation of high-level tasks in Randy's classroom.

Next, Randy actively participated in all of the ESP activities. He frequently contributed ideas and solutions to the whole-group discussions of mathematical tasks. Randy offered a few comments throughout Sessions 1 and 2 about the level of cognitive demands of the tasks. In Session 2, Randy appeared to be very engaged in considering the implementation of high-level mathematical tasks. He made 10 contributions during these discussions, with one other participant making 7 comments and the rest of the ESP teachers each making 5 comments or less. Three of Randy's comments focused on orchestrating whole-group discussions and how the

facilitator supported his learning through the use of questioning. In Session 2, Randy also reflected from The Case of Ron Castleman (Stein, et al., 2000) to his own practice, commenting that the general lessons he learned from the case were to take time to reflect on his lessons and to begin setting the expectation that his students work on high-level tasks. He also shared with the whole group an epiphany that his group had during the small group discussion, "We found in paragraph 17, what we thought was the key to the whole case, 'If I could only find a way to support students without telling them how to do it.'" (video transcript, Session 2, 11/06/04). Randy's evaluation from Session 2 indicated that he was thinking about the implementation of high-level tasks in his own classroom: **"**Implementation of high-level tasks is an issue for everyone. I was worried it was just me" (written artifacts, Session 2, 11/06/04).

At the beginning of Session 3, teachers were asked write in their journals about whether they were thinking differently about anything in their own classroom following the first two ESP sessions. Randy responded:

> After the first two sessions, I have thought about implementing more high-level tasks in the classroom. Currently, I implement some of these tasks as extra-credit assignments. With the grind of getting through the required material, it is difficult to find the time to implement new tasks into my daily lessons. In other words, I'm still working on finding ways to change my lesson plans so that they include tasks such as the ones in this program [Randy's journal, Session 3, 1/08/05].

This statement seems consistent with Randy's Winter data collection, which would have occurred in the month following Session 3. At that point, Randy had begun to incorporate high-level tasks into his classroom, but both he and the students were still becoming accustomed to implementing the tasks at a high-level.

Randy continued to be a major contributor of ideas throughout Sessions 4, 5, and 6, and his comments indicated that he was continuing to think about how to improve his practice. He commented in Session 4 that the purpose of reading narrative cases was to "help us learn from their mistakes and duplicate the things that went well," and he considered the case stories as opportunities for "getting your experience 'graded' to see how you did. It forces you to reflect on your work and see what you can do to improve" (written artifacts, Session 4, 2/05/05). In Session 5, Randy focused on improving his facilitation of whole-group discussions: "Rather than have every student show their particular overhead transparency, it may save time and ultimately be more effective to have all presenters build upon one display. At times I get bogged down in having all students present <u>everything</u> they did" (written artifacts, Session 5, 3/05/05). Randy's focus on whole-group discussions was evident in his data, as well. He conducted a whole-group discussion in the Fall that was very procedural in nature. In the Winter, his questions during the small- and whole-group discussions maintained the high-level demands, despite pressure from students. In the Spring, Randy's questions during the whole-group discussion showcased the important mathematical ideas in the lesson and allowed him to assess students' understanding of those ideas. In his final interview, Randy noted that the ESP workshop helped him "take time to stop and think about" the tasks he was using in his own classroom and "how to keep the tasks at a high-level." Randy's data appears to substantiate his self-assessment.

Randy portrays several characteristics that can be generalized to other teachers with similar patterns of change. First, Randy and five other teachers who were using reform curricula but had low task means in the Fall data collection experienced a substantial jump in task means from Fall to Winter. Second, Randy exemplifies the nature of participation in the ESP sessions of a group of eight teachers who experienced growth in both their selection and implementation of

high-level tasks. They frequently volunteered during all whole group discussions, particularly during discussions about the implementation of high-level tasks. Their written artifacts (session evaluations and journals) indicated that, at the end of each ESP session, they were thinking about the implementation of high-level tasks in their own classrooms.

4.4.4.2. A closer look at a teacher who used high-level tasks prior to ESP but improved in implementation

Nellie was teaching in an affluent suburban school during her participation in ESP, and had 4 years of teaching experience. She was recommended as a participant in ESP by her building principal and other faculty in her school. One other teacher from her school was participating in Cohort 2 along with Nellie, and two teachers had previously participated in Cohort 1 of the ESP project. Nellie submitted data for a 9th-10th grade Algebra I class using a traditional textbook. However, Nellie often supplemented lessons and entire units with self-created materials and pilot materials from reform-oriented algebra curricula. These materials were consistently high-level, even prior to Nellie's participation in ESP. Only one of the 15 main instructional tasks in her data collections was from the text, and this was an exploration from the resource materials that was rated as high-level, as well. She did, however, use the text as a source of skills review, and often assigned practice problems for homework that were at the level of procedures without connections.

Nellie entered ESP with a fairly high knowledge of the cognitive demands of mathematical tasks. Her pre-workshop task sort score of 27 was above the ESP teachers' pre-workshop mean of 24.21, and was the 6th highest of the 19 ESP teachers. Nellie's main instructional tasks had a mean of 3.2 in all three data collections. Hence, Nellie entered the ESP

project already using high-level tasks and did not exhibit change over the course of her participation in the project. In contrast, Nellie's implementation of high-level tasks did show improvement between the Fall, Winter, and Spring data collections. In the Fall, two of the three student work tasks were high-level (i.e., one scored a 4 and one scored a 3), but were both implemented as procedures without connections (i.e., a score of 2). Students used a single strategy to solve the tasks (finding a numeric pattern from a table of values), stopped working when they had reached a numerical answer, and did not connect the mathematics (exponential growth) nor their answers to the context of the problem. In one of the tasks, not one student responded to the questions that were intended to elicit higher-level thinking. Nellie noted in her data packet that none of the students' responses could be considered as examples of high-level work.

The Winter student work, which focused on modeling liner equations, was high-level in *Potential* and *Implementation*. Students' work was characterized by the use of multiple representations, and students were improving in their ability to explain their thinking on the questions that required high-level thinking. Similarly, two of the three student work tasks were implemented at a high level in the Spring data collection. One task declined into procedures without connections, and the following is Nellie's assessment of the task implementation:

> Students were to work together in groups… Many of them were using graphing calculators, and … they weren't using parentheses in the appropriate places. Instead of this lesson helping students understand application of rational expressions, it turned into a calculator entry activity. The discussion DID (*emphasis in original*) have one benefit. It helped them to understand that, in order to evaluate rational expressions, they had to make the calculator

follow the correct order of operations. I think in this process, students forgot what the numbers they were plugging and chugging actually represented.

As was the case with the student work that declined in the Fall, Nellie was very reflective about her task implementation and sensed when high-level demands had not been maintained.

Nellie's Fall observation was similar in nature to her Fall student work. The high-level task demands declined to procedures without connections, as students did not complete the parts of the task that required high-level thinking. Nellie pressed them to answer these questions as she circulated amongst the groups, but the focus seemed to shift to finding numerical answers rather than determining patterns that could be generalized into the rules for simplifying exponents. Nellie appeared to have developed a set list of questions that she repeatedly asked each group. The whole-group discussion of the task was characterized by students providing answers, followed by Nellie evaluating their answers as correct or incorrect and asking another student to explain the procedure. A portion of the task entitled "Drawing Conclusions" that could have engaged students in high-level thinking had not been completed by students and was not discussed during the lesson.

The Winter and Spring observations were different in character. Both featured tasks that were maintained at a high-level, and students were held accountable for addressing the aspects of the task that required high-level thinking and reasoning. For example, during the winter observation, when students were not making an important connection between the equation and the graph of a linear function, Nellie required students to "turn and talk to a partner" and explain what they thought was happening in the graph. She circulated and listened to each pair, sometimes prompting them with questions that were tailored to their thinking (rather than a set list of questions for every pair as observed in the Fall). After about ten minutes, she repeated the

original question to the whole group and specifically targeted a few students to share their ideas. Nellie noted in the post-lesson interview that she had picked up the "turn-and-talk" move in the classroom of her mentor teacher when she was conducting her preservice teaching. She attributed much of her teaching style to her preservice MAT program and to her mentor, who was recognized as a mathematics teacher-leader in the region.

On the evaluation of the first ESP session, Nellie remarked, "I am on track with ESP's philosophy of teaching" (written artifacts, Session 1, 10/2/04). Her evaluations of the session that followed indicated that she was already selecting high-level tasks for instruction (i.e., "How to *keep* high level tasks as high level tasks" [written artifacts, Session 2, 11/06/04; emphasis added]) and was already incorporating many of the factors for maintaining high-level cognitive demands (i.e., "*Keep* pressing for justification" [written artifacts, Session 5, 3/05/05; emphasis added]). Nellie still appeared to benefit from her participation in ESP in several ways. Her post-workshop task sort score increased to 35, exceeding the post-workshop task sort mean of 28.79, and placing her as the second highest score of the 19 ESP teachers. Hence, Nellie improved her knowledge of the cognitive demands of mathematical task following her participation in the ESP workshop. More importantly, the ESP workshop appeared to have helped Nellie think about aspects of her teaching in ways that catalyzed improvements in her implementation of high-level tasks. Nellie made frequent contributions in each session, and volunteered at least 4 ideas per session that pertained to the implementation of high-level tasks. When asked in Session 2 whether the Mathematical Tasks Framework resonated with her experiences in the classroom, she responded in her journal:

> This makes complete sense to me, because when I attempt to implement any task, I always put it on myself if something goes bad. In order to be successful and have student learning

> happen, I have to set up the task at a high level and my students must be expected to perform at a certain level without frustration (Nellie's journal, Session 2, 11/06/04).

In Session 3, when asked what the first two sessions caused her to think about, she indicated: "I want students to maintain high-level work during lessons. I have paid attention to more of the details in my lessons: wait time, questioning, types of student responses." In Nellie's self-assessment of her learning in the ESP project, she identified the use of questioning as an area of improvement. She noted that thinking about questioning ahead of time allowed her to better identify the mathematics she wanted students to learn, press them for those ideas during both small- and whole-group discussions, evaluate their understanding, and "if they do not respond, then they are forced to talk in their groups" (Nellie's journal, Session 6, 5/07/05). Nellie's improvement in these areas is evident in the data collection. She begins to hold students accountable for engaging with the high-level aspects of instructional tasks, improved her use of questioning, and utilized the "turn-and-talk" strategy when students did not provide responses to indicate their understanding of the main mathematical ideas.

In the final interview, Nellie explains that though she had been exposed to similar ideas in her MAT program and in other professional development experiences, the ESP workshop had refreshed and strengthened her ability to "teach the way I want to teach." Specifically, she noted:

> It was a continuation of my MAT program. But in four years, I did lose a few things along the way, and I found myself becoming a little too traditional, giving answers. …And I found myself being aware of when I was doing that, … and this has helped me reflect on what I was doing and what students were taking away from the lesson. And throughout this school year, I've become more reflective, and really thinking about the questions that I ask and how that impacts students learning. … (ESP) put me back on track of where I needed to go. I'd sit in

> conferences… and come out saying 'oh I feel refreshed' …but once you go back in your classroom, its the real world again and you just naturally start to fall back into that. Its hard, its not easy to stay where I want to stay. And now I can look harder at myself, and I think the reflection, thinking of what I am doing rather than just doing it, is a benefit of the ESP program" (final interview, audiotape transcript, 6/24/06).

When pressed as to how the ESP workshop was different from her other professional development experiences, she stated that ESP provided a tool for planning lessons around high-level tasks (i.e., the TTLP) and a structure for reflecting on the set-up and implementation of the task (i.e., the MTF). Data on Nellie indicates that she was reflective even prior to her participation in ESP, but perhaps ESP provided a direction for her reflections that, in Nellie's own words, "put the focus back on students' learning."

Beyond being an interesting individual case, Nellie also exemplifies the characteristics of a group of five teachers who entered the ESP project consistently using high-level tasks as the main instructional tasks in their classrooms, but improved in their ability to maintain high-level demands during implementation. This group of teachers described ESP as supporting and refining their current instructional practices. Consistent with Nellie, their comments indicated that the ideas in ESP were not new or revolutionary information and often centered on <u>improving</u> implementation; for example [emphasis added]: "*Continue* to keep high standards for students (Madeline, Session 1, 11/05/04); "Pay *closer* attention to the order in which I have students present solutions/strategies to the class to help facilitate connections" (Michelle, Session 2, 11/05/04); "give myself *more* time to consider implementation" (Kelsey, Session 2, 11/05/04). This group of teachers was already convinced of the value of high-level tasks, had already experienced the trials and tribulations of maintaining high-level task demands during instruction.

They were ready to begin the process of improving their ability to implement high-level tasks. Randy's group had to first consider how to use a high-level task as a main instructional task before they could consider issues of implementation. As such, scores from Randy's group increase for task selection between Fall and Winter, and they exhibit improvements in task implementation between Winter and Spring.

4.4.4.3. A closer look at a teacher who did not exhibit improvements in instructional practices

Cara graduated from the same pre-service teacher program in the same year as Nellie. Both Cara and Nellie taught in affluent suburban schools, though in different school districts, both had been teaching for 4 years, and both submitted data collections from Algebra I classes for 9-10th graders. Both teachers were using a traditional algebra textbook, which they supplemented with self-created materials and materials from other resources. And finally, both teachers created materials that were similar in appearance, such as in the use of clip art or web graphics and the use of problems with multiple parts.

Despite this list of similarities, Cara's and Nellie's instructional practices and the nature of their participation in the ESP workshop were quite different. Cara's pre-workshop task sort score of 19 was well below the group mean of 24.21, and placed her as the 4th lowest score of the 19 ESP teachers. Cara was teaching the algebra class on a block schedule, so her Winter data collection represented the beginning of the algebra course and her Spring data collection represented the end of the course. Her task means for both Winter and Spring were a 2.0, with all tasks at the procedures without connections level. Cara's main instructional tasks consisted of self-created materials very impressive in appearance, but still low in their level of cognitive

demand. She utilized internet resources and incorporated real-world examples, but over-structured the mathematics to the point that students only needed to perform the computations or procedural aspects of the algebra. Cara did not elect to participate in classroom observations, so data about her instructional practices is described based on her collections of student work.

Cara's student-work tasks and implementation in Winter and Spring were rated entirely as procedures without connections (i.e., a score of 2). For example, one set of student work from the Winter data collection consisted of students' responses to a handout that structured a lesson on the distributive property. Cara's lesson handout was self-created and included clip art, but did not include high-level tasks. Directions to students consisted of, "Simplify the following expressions. Remember to combine like terms if possible" (Cara, Winter student work, 2/05/05). This lesson occurred in the month between Sessions 3 and 4 of the ESP workshop (i.e., between January 8, 2005 and February 5, 2005). Interestingly, in Session 3, teachers engaged with a task and read a narrative case featuring the use of area models and algebra tiles to model multiplication with algebraic terms (i.e., monomials and binomials). Ways of visually illustrating the distributive property with grid paper and with algebra tiles, and methods for connecting multiplication of algebraic terms to students' prior knowledge of multiplication with whole numbers, was explicitly discussed throughout the session. In the month following Session 3, two days of instruction in Cara's classroom focused on the distributive property. None of the tasks or student work used for instruction utilized area models, grid paper, algebra tiles, or other opportunities for students to make sense of the procedures for multiplying algebraic terms, even though that specific mathematical content had been explored in the previous ESP session. Instead, students' written work consisted of procedures and answers. Similarly, student work in the Spring data collection focused on factoring algebraic expressions. Lessons were structured

through self-created handouts, where students were given three methods of factoring polynomials and were then provided with opportunities to practice using each of the methods over the five days of instruction represented in the data collection. Directions consisted of, "Using the tools from your toolbox, factor the following polynomial as much as possible" and "Factor each of the following polynomials using the 3 steps of factoring that you have learned about so far" (Cara, Spring student work, 5/07/05). Hence, Cara's instructional practices in the Spring were characterized by the selection and implementation of procedural tasks with no connections to meaning or understanding. Students in her classroom engaged in following rote procedures to produce correct answers, in contrast to students in Randy's and Nellie's classrooms, in which observed instruction and student work provided evidence of students' opportunities to engage in high-level thinking and reasoning.

Cara's participation in ESP was also quite different than Randy and Nellie. Cara made two or less contributions per session, and these contributions typically occurred during discussions of mathematical tasks. Cara made three 3 comments about implementation throughout the course of her participation in ESP, whereas Randy and Nellie made several comments per session. There was also a difference in the nature of Cara's comments on the evaluation sheets. In Session 2, she stated her intention to "convert some of my pre-existing tasks into higher-level tasks" (written artifacts, Session 2, 11/05/04). Cara wrote in her journal in Session 3 that she was "trying to find more activities/tasks that allow students time to explore and come to their own conclusions" and that she had "looked more closely at making some of my tasks higher level tasks" (Cara's journal, Session 3, 1/08/05). Cara requested resources for high-level tasks in Sessions 2, 3, 4, and 6; and she commented in Sessions 3, 4, and 6 that she intended to use the high-level tasks presented in the sessions and by colleagues in ESP. Even in

Session 6, Cara appeared to still be thinking about selecting high-level tasks to use for instruction in her classroom, which she had been considering since Session 2. Cara did improve in her knowledge of the cognitive demands of mathematical tasks. Her post-workshop task sort score increased from 19 to 28, slightly lower than the ESP teachers' average of 28.79 but placing her at the median of the post-workshop task sort scores. However, as evidenced in her data collection, the increased knowledge of the cognitive demands of mathematical tasks did not prompt Cara to select high-level tasks for instruction in her classroom. Furthermore, the lack of contributions regarding implementation, coupled with the lack of improvement in her implementation data, seem to indicate that Cara was not deeply engaged in considering the implementation of high-level tasks within the professional development or within her own classroom.

Cara's participation in the ESP sessions typifies the participation of teachers who exhibited no change (two teachers) or exhibited change in task selection but not task implementation (three teachers). This group of teachers rarely volunteered during the sessions, and most of their contributions occurred during discussions of mathematical tasks. As a group, they volunteered far less frequently than teachers who exhibited change in task implementation, and rarely offered comments on the implementation of high-level tasks. While their engagement in the small group discussions was not documented, their lack of contributions regarding implementation during the whole-group discussions and in written artifacts appears to support the evidence in the data collection that they were not in the process of improving their implementation of high-level tasks, and only a subset were improving the selection of high-level instructional tasks in their own classrooms.

In considering instructional change along a continuum from task selection to task implementation, 16 of the 18 teachers in the study advanced their position on this continuum

throughout their participation in the ESP workshop. Chapter 5 will discuss the conclusions that can be drawn from the results of the statistical analyses, descriptive data, and case studies presented in this chapter.

5. CHAPTER 5: DISCUSSION

In this chapter, a discussion of what can be learned from this investigation and how the findings can inform mathematics teachers' learning and professional development more broadly is presented. The chapter begins by describing the importance of this study. Next, possible explanations for the effectiveness of the ESP professional development workshop and for the results obtained in this study are presented. This is followed by a discussion of the ways in which this investigation contributes to the knowledge base of professional development research. The chapter closes with concluding remarks and suggestions for future research.

5.1. Importance of this Study: Improving Mathematics Teaching to Improve Mathematics Learning

The results of this study are important to mathematics professional development, teaching, and learning because teachers' experiences in the ESP professional development workshop appear to have enabled them to enact instructional change in ways that have the potential to increase opportunities for students' learning. This study contributes to a small body of professional development research for which classroom observations and artifacts provide evidence of enhancements in teachers' knowledge and instructional practices following teachers' participation in the professional development intervention. ESP facilitators designed and

presented a series learning experiences that allowed teachers to focus on transforming a specific aspect of their knowledge and instructional practices – the selection and implementation of cognitively challenging tasks.

In Chapter 1, the argument was waged that improvements in the learning of mathematics would need to be preceded by improvements in the teaching of mathematics (USDE, 2000), and that professional development experiences for teachers could catalyze instructional improvements that could potentially yield increased opportunities for students' learning. Research has identified engaging with mathematical tasks that foster understanding and provide opportunities for high-level thinking and reasoning as one instructional improvement that may increase students' mathematical proficiency (Stein & Lane, 1996; Hiebert & Wearne, 1993). Hence, a promising approach to improving student achievement in mathematics would be to improve the quality of the instructional tasks in which students have the opportunity to engage and the implementation of these tasks in teachers' classrooms.

The purpose of this study was to determine whether the ESP professional development workshop influenced teachers' knowledge, selection, and implementation of mathematical tasks. This phenomenon was studied by conducting a pre- and post-workshop assessment of teachers' knowledge of the cognitive demands of mathematical tasks, collecting artifacts and observations from the teachers' classrooms throughout their participation in the workshop, and collecting data from the professional development sessions.

The results of this study provide evidence that, following their participation in the ESP workshop, teachers demonstrated growth in their knowledge, selection, and implementation of cognitively challenging mathematical tasks. Table 5.1 summarizes the results obtained from this study. From Fall to Spring, ESP teachers significantly increased: 1) their knowledge of the

Table 5.1.
Summary of Results

Data Source	Significant Increase from Fall to Spring	Significantly Higher than Contrast Group
Task Sort Scores	Yes	Yes
Task Collection		
Mean Task *Potential* Scores	Yes	NA
Number of High-Level Tasks	Yes	NA
Student Work		
Mean Implementation Scores	Yes	NA
Number of High-Level Implementations	Yes	NA
Observations		
Mean Implementation Scores	No	Yes

cognitive demands of mathematical tasks (i.e., their task sort scores); 2) the level of cognitive demand of the main instructional tasks used in their classrooms (i.e., the mean task *Potential* score and the number of high-level task used as the main instructional tasks); and 3) their ability to maintain high-level cognitive demands during instruction (i.e., the mean *Implementation* score for student work tasks and the number of high-level implementations in student-work tasks and lesson observations). The mean *Implementation* score for lesson observations did not increase significantly (partially due to the small size of the sample). However, in six teachers' classrooms, the amount of increase in lesson *Implementation* scores and the score levels between which the increase occurred were important for moving task implementation from low- to high-level cognitive demands. As compared to a contrast group of secondary mathematics teachers who did not participate in the ESP professional development workshop, the ESP teachers had

significantly higher 1) task sort scores, and 2) task and implementation mean scores for classroom observations. Hence, ESP teachers' knowledge and instructional practices changed significantly following their participation in the ESP workshop and were significantly different from a group of teachers who did not participate in the workshop.

Most critically, ESP teachers improved their knowledge and instructional practices along a dimension of teaching linked by prior research to increased student achievement in mathematics. Although this study did not measure the impact of the changes in teachers' knowledge or instructional practices on students' learning, evidence from other studies (e.g., USDE-NCES, 2003; Stein & Lane, 1996; Hiebert & Wearne, 1993) suggests that increased exposure to cognitively challenging tasks and extended engagement with high-level cognitive demands increases students' learning of mathematics. Furthermore, research by Hill, Rowan, and Ball (Hill, Rowan, & Ball, 2005; Hill & Ball, 2004) has established a link between teachers' knowledge of mathematics for teaching and student learning outcomes. Following their participation in the ESP workshop, teachers who participated in this study increased their knowledge of the cognitive demands of mathematical tasks, more frequently selected high-level tasks as the main instructional tasks in their classrooms, and improved their ability to maintain high-level demands during implementation. Hence, based on evidence from prior research, the changes in teachers' knowledge and instructional practices identified in this study hold potential for improving students' opportunities for learning.

The ESP workshop appears to have provided teachers with opportunities for transformative learning (Thompson & Zeuli, 1999), in ways that, as called for by *Before Its Too Late* (USDE, 2000), could ameliorate the deficiencies in students' opportunities to learn mathematics with understanding. The following section will posit an explanation for the

effectiveness of the ESP workshop in promoting teachers' learning and instructional change as identified by the results of this investigation.

5.2. Explanations for the Results

This section offers explanations for the results obtained in this study. How did the ESP workshop promote teachers' learning and instructional change? Did teachers simply provide the types of responses and tasks that were modeled in the ESP workshop? Alternatively, did teachers change their prior conceptions of effective mathematics teaching and learning, and begin to change their instructional practices in accordance with these new ideas? The merits of these conflicting explanations will be explored in the remainder of this section.

5.2.1. The ESP Workshop Provided Transformative Learning Experiences for Teachers

The ESP professional development workshop provided learning opportunities for teachers that appeared to transform their knowledge, selection, and implementation of cognitively challenging mathematical tasks. This raises questions regarding the aspects of the ESP workshop that were effective in promoting teachers' learning and instructional change.

The ESP professional development workshop was designed and implemented in ways consistent with social-constructivist theories of teacher-learning (Simon, et al., 2000; Cobb, Yackel, & Wood, 1991; Simon & Shifter, 1991). Three components of social-constructivism hold implications for the design and facilitation of learning experiences for teachers: 1) build on teachers' prior knowledge and beliefs; 2) allow teachers to wrestle with new conceptions of teaching and learning mathematics that conflict with their prior knowledge and beliefs; and 3) enable social interactions that stimulate, sustain, and support teachers' consideration of new

and/or conflicting ideas. The ESP workshop embodied all three components. First, ESP teachers' learning was supported by the continued refinement of prior knowledge and the gradual application of new knowledge to their own instructional practices. The initial activity in Session 1 of the ESP workshop engaged teachers in solving two mathematical tasks with contrasting levels of cognitive demand. Teachers were asked to compare the tasks by reflecting on their own experiences, and this discussion generated several ideas that were built upon during the ensuing discussion of the levels of cognitive demand of mathematical tasks. This sequence of activities provided an initial, shared experience accessible to all teachers in the group and provided the foundation for deeper consideration of the cognitive demands of mathematical tasks. Similarly, teachers considered the implementation of cognitively challenging tasks in narrative cases only after solving the tasks as learners and reflecting on what supported their own high-level engagement. Teachers' "case stories" (Ackerman, Maslin-Ostrowski, & Christensen, 1996) of implementing cognitively challenging tasks in their own classrooms progressed from a general description in Session 3 (i.e., a written account of what factors influenced the lesson) to a more detailed representation of instruction and learning in Sessions 5 and 6 (i.e., bringing in student work or questions they asked during the lesson).

Second, written artifacts and videotapes from the sessions provide evidence that ESP teachers wrestled with new ideas that conflicted with prior conceptions. In Session 1, for example, a discussion arose within the large group about the nuances of the word "explain" as an indicator of high-level cognitive demands. The group consensus was that a task containing a prompt to "explain" would have high-level cognitive demands. Five participants contributed their ideas in support of this contention, stating that explaining the reasoning or providing justification for mathematical work invoked high-level cognitive processes. Three participants

then highlighted the difference between *explaining how* versus *explaining why* and that *explaining the steps to a procedure* would involve low-level cognitive demands. When pressed by the facilitator to consider whether the presence of the word "explain" was a sufficient criterion for high-level demands, the majority of participants appeared to have reconsidered or refined their initial position and responded, "No." One participant, however, was still not convinced. Dana turned to other members of her small group and questioned, "No? But if they generate an explanation, that's high-level. That's what made all the other ones high-level, right?" [video-transcript, Session 1, 10/02/04]. Two other participants offered ideas to the whole group, and a member of Dana's small group responded directly to Dana (inaudible to the video camera). At that point, Dana nodded in agreement and said aloud to her group, "So *explain* doesn't necessarily mean high-level." This example illustrates the whole group and an individual participant refining their prior conceptions about the cognitive demands of mathematical tasks. Other instances where participants are wrestling with or accommodating new ideas were also evident in the data. Teachers who improved their task selection and implementation (as represented by Randy) publicly shared several epiphanies that occurred within the ESP sessions or within their classrooms as they reflected on ideas from the ESP sessions (e.g., high-level tasks were only used as extra credit; the student work all looked the same; building a conceptual understanding *before* using short-cuts and procedures).

Several design features of the ESP sessions supported social interactions. Small groups were often configured to contain individuals with different views on effective mathematics teaching and learning. In this way, we increased the likelihood that differing viewpoints would arise within the small group discussions and teachers would need to consider and respond to ideas about teaching and learning mathematics that conflicted with their own. Groups were

sometimes structured to encourage collaboration among teachers from the same school, who taught in similar grade-levels (middle school vs. high school), or who taught similar courses (i.e., algebra, geometry, pre-calculus, etc.). This was intended to tighten the connections to teachers' own classrooms, as they shared tasks and ideas about effective implementation with other teachers in similar instructional settings. At other times, groups were purposefully constructed to associate teachers from different schools, grades, or course assignments. This arrangement brought different mathematical expertise and problem-solving strategies into the discussions of mathematical tasks and students' mathematical thinking.

The nature of the tasks posed to teachers also contributed to the richness of the small- and whole-group interactions. The practice-based professional learning tasks (Smith, 2001; Ball & Cohen, 1999) focused on understanding important mathematics, analyzing student thinking, and reflecting on mathematics pedagogy - "activities that are at the heart of a teachers' daily work" (Smith, 2001, p.2). Within each activity, questions were crafted by the ESP team that allowed all teachers to enter the discussion, that prompted teachers' thinking about issues associated with selecting and implementing high-level tasks, and that promoted generalizations from the specific activity to the teaching and learning of mathematics more broadly. In this way, the activities used within the ESP sessions were "group-worthy" tasks as described by Lotan (2003).

Finally, whole-group discussions throughout the ESP sessions were characterized by facilitation moves that initiated, maintained, and supported teachers' consideration of critical ideas. The ESP facilitators consistently used moves such as encouraging the extended discussion of a topic (i.e., "Any other ideas about…?" or "Say more about .."), redirecting questions to the group, revoicing participants' comments and ideas, and clarifying the poles of potential debates (i.e., "Let's take two positions. Iris is saying doing mathematics and Madeline is saying

procedures with connections." [video transcript, Session 4, 2/05/05]). These facilitation moves have been identified as aspects of discourse that promote learning with understanding for teachers (Boston, 2003; Remillard & Geist, 2002; Barnett, 1991; Cobb, Yackel, & Wood, 1991) as well as for students (O'Connor & Michaels, 1996; Forman, Larreamendy-Joerns, Stein, & Brown, 1998).

Hence, the ESP workshop provided learning experiences that built on prior knowledge, allowed teachers to wrestle with new ideas, and provided social interaction that supported and enhanced teachers' consideration of new ideas. Increases in teachers' ability to select and implement high-level tasks suggest that teachers not only wrestled with new ideas about the level, selection and implementation of cognitively challenging tasks, but accommodated these new ideas in ways that influenced their knowledge and instructional practices. Explanations for the increases in teachers' knowledge and instructional practices will be waged in the sections that follow.

5.2.2. Increasing Teachers' Awareness of the Influence of Cognitively Challenging Tasks on Students' Learning

One explanation for the significant increases in teachers' task sort scores and selection of high-level tasks is that teachers increased their awareness of the influence of cognitively challenging tasks on students' learning of mathematics. ESP teachers' task sort scores increased from pre- to post-workshop, and were significantly higher following their participation in ESP than the scores of a group of contrast teachers who did not participate in the ESP workshop. The nature of the improvements in teachers' pre- to post-workshop task sort responses provide evidence that teachers did not simply learn the "correct" answers throughout their participation

in ESP. Teachers still classified similar numbers and categories of tasks incorrectly; the improvements occurred in teachers' *criteria* and *rationales* for high- and low-level tasks. This finding is consistent with research by Stein and colleagues (Stein, Baxter, & Leinhardt, 1989), where a notable difference existed between the *type of criteria* provided by the novice and the experts for their categories on the task sort. Criteria provided by the experts were less focused on surface-level features and more apt to describe a connected, rich understanding of functions. ESP teachers' pre-workshop task sort criteria and the criteria identified by the contrast group reflect characteristics of the novice; attention to superficial features of tasks that were largely irrelevant to the level of cognitive demand (i.e., a task is high-level because of the word 'explain,' because the task is a 'word-problem,' or because the task is perceived to be 'difficult;' a task is low-level because the task contains a diagram). ESP teachers' post-workshop task sort responses illustrate less focus on superficial criteria and an increased focus on features and characteristics of the level of cognitive demands of mathematical tasks. In this sense, ESP teachers' post-workshop task sort criteria became more "expert-like" and reflected an enhanced knowledge of the characteristics of mathematical tasks that influence students' opportunities for high-level thinking and reasoning.

The emergent criteria in teachers' post-workshop task sort responses provide further evidence that teachers became more aware of how high-level tasks support students' learning. The most striking characteristic about the nature of ESP teachers' criteria for high- and low-level tasks on the post-workshop task sort was the close connection between the emergent criteria and the topics publicly discussed during the ESP sessions (see Figure 4.2). Both the TAG (Figure 2.1) and teachers' experiences within the ESP workshop appear to have influenced their thinking about the cognitive demands of mathematical tasks. The fact that teachers incorporated a greater

number of high-level tasks as the main instructional tasks in their classrooms suggests that the TAG also served as tool for thinking about the level of tasks in their own classrooms. In their research on teachers' instructional change following participation in study groups focused on the cognitive demands of mathematical tasks, Arbaugh & Brown (2002) argue that the TAG provided a framework through which teachers in their study learned to critically examine the tasks they selected for instruction in their own classrooms.

The argument that teachers' curricula dictated the instructional tasks teachers used in their classrooms was nullified by ANOVA results indicating that the use of a reform vs. traditional curricula was not a significant influence on their knowledge, selection, or implementation of high-level tasks. This result is particularly surprising with regard to task selection; reform curricula are specifically designed to contain a greater percentage of cognitively challenging tasks (AAAS, 2000; USDE, 1999). The non-significant results do not indicate a lack of high-level tasks in reform curricula. Rather, they reflect a prevalent research finding that mathematics teachers do not always use reform-oriented curricular materials as intended by the curriculum developer (Remillard & Bryans, 2004; Remillard, 1999; Lloyd, 1999; Lloyd & Wilson, 1998). Teachers' conceptions "act as critical filters" (Lloyd & Wilson, 1998, p. 250) that govern their use of curricula in ways that can be supportive of or antithetical to reform-oriented mathematics pedagogy. In the current investigation, this statement was true of teachers using both reform and traditional curricula. Close examination of the main instructional tasks revealed that teachers using reform curricula often did not use the high-level tasks offered by the curricula as their main instructional tasks, and teachers using traditional curricula often used supplementary materials or created extensions that increased the cognitive demands of the tasks in their curricular materials. Both groups selected the main instructional tasks for their data

collections in all of the ways identified by Remillard & Bryans (2004, p. 363): *guided by* the curricula (teachers used their reform or traditional curricula as the source of the task and of the lesson activities); *drawn from* the curricula (teachers used the tasks from the curricula, but implement these tasks in their own way), *adapted from* the curricula (teachers altered the tasks in the curricula), or replaced the curricula with *other resources* or tasks of their *own-design*. Remillard & Bryans refer to this as teachers' "orientation toward curricula," defined as "a set of perspectives and dispositions about mathematics, teaching, learning, and curriculum that together influence how a teacher engages and interacts with a particular set of curricular materials…" (2004, p. 364). ESP teachers increased their knowledge of the cognitive demands of mathematical tasks, and as argued earlier, increased their awareness of how high-level tasks support students' learning. Hence, by enhancing teachers' knowledge of the cognitive demands of mathematical tasks, teachers changed their orientation toward their curricula (reform <u>or</u> traditional) in ways that supported the selection of high-level instructional tasks in their own classrooms.

ESP teachers were presented with research (e.g., Stein & Lane, 1996) and narrative cases illustrating the influence of high-level tasks on students' learning. Based on the premise that teachers will act according to their own conceptions of what is best for their students (Remillard, 1999; Borko & Putnam, 1995; Thompson, 1992), a heightened awareness of the value of high-level tasks in supporting students' learning may have prompted ESP teachers to use a greater number of high-level tasks for instruction in their own classrooms.

5.2.3. ESP Teachers Increased their Selection of High-Level Instructional Tasks

In the Spring data collection, ESP teachers used significantly more high-level tasks as their main instructional tasks than in the Fall data collection. In what ways does this difference in the main instructional tasks submitted in teachers' data collections generalize to the main instructional tasks used in teachers' classrooms on a regular basis? One possibility is that the results do not generalize to teachers' everyday instructional practices; rather, teachers inferred from the professional development activities that we were "looking for" high-level tasks in their data collections, and used such tasks as the main instructional task during the week of data collection only.

While this scenario would still provide evidence of teachers' knowledge of the cognitive demands of mathematical tasks (i.e., teachers would need to be able to recognize tasks with high-level cognitive demands in order to specifically select these tasks for use in their data collection), it does not account for improvements in teachers' ability to *implement* high-level tasks in ways that maintained high-level cognitive demands. Implementing a high-level task in ways that maintain the cognitive demands is not a trivial endeavor, as document by large-scale studies such as QUASAR (Henningsen & Stein, 1997), the TIMSS 1999 Video Study (USDE-NCES, 2003), and Horizon Research, Inc. (Weiss & Pasley, 2004; Weiss, et al., 2003). If teachers were not attempting to implement high-level tasks on a regular basis, no significant improvement would be evident in their task implementation data. Teachers' comments and written reflections from the professional development sessions indicate that they were using high-level tasks on a consistent, on-going basis. Hence, the triangulation of improvements in task selection, improvements in task implementation, and teachers' self-reports provides consistent evidence

that, throughout their participation in ESP, teachers were increasingly selecting high-level tasks as the main instructional tasks in their classrooms.

Another argument might be that, despite statistical significance in the increase in task means between Fall and Spring, was an increase of 0.47 really important in terms of teachers' use of instructional tasks or students' opportunities for learning? An increase between score levels 1 and 2, with the Spring task mean at or barely exceeding a 2.0, would have indicated that teachers were still using low-level tasks following their participation in ESP, just different types of low-level task (more 2s than 1s rather than vice versa) than in the Fall. Similarly, an increase of half a point between score levels 3 and 4 would have indicated that teachers enhanced the high-level tasks that they were already using prior to their participation in ESP. In both scenarios, the number of high-level tasks in each data collections would not have increased, and teachers' use of high-level instructional tasks would not have changed in ways desired by this study. However, the increase in task means occurred exactly where it was crucial for influencing the number of high-level vs. low-level tasks used in teachers' classrooms, moving teachers' main instructional tasks from predominantly low-level (i.e., a score of 1 or 2) to predominantly high-level (i.e., a score of 3 or 4) following their participation in ESP. This was evidenced by significant increases in task means and in the number of high-level tasks between Fall and Spring.

Following their participation in ESP, teachers were more frequently selecting high-level tasks as the main instructional tasks in their own classrooms, thereby increasing the likelihood of a greater number of high-level implementations. While this is true, improvements in the student work implementation were not merely the result of teachers using better tasks. The following

section will describe how improvements in implementation indicate that ESP teachers' were in the process of instructional change.

5.2.4. ESP Teachers were in the Process of Instructional Change

Comparisons of the implementation of high-level student work tasks indicated that high-level demands were less likely to decline in Spring than in Fall. As noted earlier in this chapter, the difficulty of maintaining high-level task demands is well-documented in research (Weiss & Palsey, 2004; USDE-NCES, 2003; Henningsen & Stein, 1997). This fact serves to dissipate the claim that improvements in implementation were evident only in the student work submitted in the data collections. If teachers could simply will themselves to implement tasks at a high-level, we would expect every high-level task in the data collection to have been maintained at a high-level during implementation. Instead, patterns in the data, teachers' self reports, and comments and written artifacts from the professional development sessions indicate that ESP teachers were working toward improving their ability to maintain high-level cognitive demands during instruction. Teachers were in the process of instructional change, and evidence of improvement is an indication that they were making an effort to implement and maintain high-level tasks in their classrooms at other times beyond the weeks of data collection. *Working toward* instructional change does not imply that ESP teachers were experts at implementing high-level tasks following their participation in ESP. Rather, ESP teachers were improving their implementation of high-level tasks, and these improvements pushed implementation means over the demarcation line between high- vs. low-level implementation (i.e., between score levels 2 and 3). Once again, the increases occurred exactly where it mattered most in term of students' opportunities to engage with predominantly high-level tasks vs. predominantly low-level tasks.

Instruction characterized by low-level cognitive demands in the Fall data collection evolved into instruction characterized by high-level cognitive demands in the Spring.

ESP teachers engaged in several opportunities to discuss the implementation of cognitively challenging tasks throughout the ESP workshop, as related to professional development materials and to their own classrooms (see Table 4.14). Evidence from the sessions (i.e., comments, case stories, and classroom artifacts) indicates that several teachers were thinking about issues of implementation, and the data indicated that those same teachers were improving the implementation of high-level tasks in their own classrooms. What explanation can be waged for the lack of improvement in some teachers' instructional practices? The data did not portray any patterns in teachers' age, years of teaching, school, or school demographics; as illustrated in case studies, teachers similar along several dimensions exhibited different patterns of change. Farmer and colleagues (2003), as expressed in their levels of appropriation, recognized that individual teachers engaged in the same professional development experiences will benefit differently from those experiences. In the group of five teachers represented by Cara, two teachers exhibited no improvement in task selection or implementation and the three others improved in selection only. The most striking characteristic that appeared to separate and define these teachers as a group and differentiate them from teachers who exhibited change in implementation was the nature of teachers' verbal contributions during the professional development sessions. Teachers who did not improve contributed far less frequently, and rarely contributed ideas about implementation. Comments and reflections from teachers who exhibited improvements indicate that they were considering issues of implementation and making connections between the professional development experiences and their own classrooms.

Perhaps the group of teachers who exhibited little instructional change was not as easily convinced of the value of high-level tasks in supporting students' learning. Perhaps they experienced greater difficultly in incorporating high-level tasks into their curriculum. Recall that, for ESP teachers overall, firm conclusions can be drawn that changes in knowledge and instructional practices were not the result of the type of curricula used in their classrooms. Interestingly, however, one pattern of interest was that four of the five teachers who exhibited little change (i.e., as represented by Cara) were using traditional curricula in their classrooms. Perhaps this group of teachers lacked access to resources that contained high-level instructional tasks, or felt pressure or obligation to use the tasks in their textbook, and thus enhanced their task selection at a slower pace (or not at all) compared to the other teachers. If the changes desired by this study lay along a continuum of instructional change (Smith, 1995) from using high-level tasks, to improving the implementation of high-level tasks, to implementing high-level tasks in ways that maintain the cognitive demands, ESP teachers were at different points along this continuum at the end of the workshop. Overall, 16 of the 18 teachers progressed forward from their original position. The sustained, specific focus on cognitive demands of mathematical tasks was enough to move the group of teachers forward in their selection and implementation of high-level task, but perhaps for some individuals, more time and/or more direct connections to their own practice were needed to experience significant instructional change.

5.2.5. Effectiveness of the ESP Workshop

The effectiveness of the ESP professional development workshop in moving teachers' along the continuum of instructional change is noteworthy given that the six ESP sessions consisted of 30 total contact hours with teachers. The six sessions were spread throughout the

course of a school year, so perhaps the duration of the project contributed to its success. In addition to the number of actual contact hours, teachers were given assignments that closely connected the ideas from the professional development sessions to their own classroom practices. Furthermore, the nature of the ESP activities provided frequent opportunities for several types of reflection: 1) reflection on ideas about instruction and learning as featured in the professional development materials; 2) reflection from the context of the professional development materials to instruction and learning in teachers' own classrooms; and 3) reflection on teachers' own instructional practices using ideas and frameworks provided by the professional development workshop. Close alignment between the goals of the professional development activities with the goals for teachers' instructional change (i.e., the selection and implementation of high-level mathematical tasks) created opportunities for teachers to reflect from the specifics of the professional development activities to their own classroom (Wallen &Williams, 2000; Barnett, 1998). The value of teacher reflections on instructional change has been noted throughout research and theories of effective professional development (Wallen & Williams, 2000; Smith, 2000; Ball & Cohen, 1999; Thompson & Zeuli, 1999). In this way, though the ESP workshop consisted of 30 hours of professional development, the teachers were engaged with the ideas and tools from ESP beyond the constraints of the time spent in the actual workshop.

Significant increases in teachers' knowledge and instructional practices support the contention that teachers resonated with the frameworks and tools provided by the ESP workshop in ways that allowed ideas from the professional development to travel into teachers' classrooms (Smith, Boston, & Steele, 2006). Examples of "tools" provided to teachers throughout the ESP workshop include the Task Analysis Guide (TAG; see Figure 2.1), the Mathematical Tasks Framework (MTF; see Figure 2.2), the factors that influence the maintenance and decline of

high-level cognitive demands (see Figure 2.3), and the "Thinking Through a Lesson" protocol (TTLP) (Hughes & Smith, 2004). Ideas from the TAG permeated teachers' post-workshop task sort criteria, indicating that the TAG provided a structure and/or a language for teachers to describe the cognitive demands of mathematical tasks that they did not have access to prior to their participation in the ESP workshop. For the other tools provided to teachers, only self-report data exists on the extent to which teachers came to view their own practice through these frameworks. For example, teachers' journal entries indicated that the MTF resonated with their own experiences in implementing high-level tasks; teachers' comments and written artifacts for the case stories show that teachers utilized the MTF and the factors to reflect on their own teaching; and Nellie's interview indicated that she found the TTAL useful for considering how to maintain high-level cognitive demands while planning a lesson. From teachers' self-reports, comments and written artifacts from the ESP workshop, and case stories of their own trials and tribulations in implementing high-level task, the ESP team has speculated that the tools enabled teachers to generalize the ideas about selecting and implementing high-level tasks explored during the ESP sessions and apply those ideas to their own instructional practices (Smith, Boston, & Steele, 2006). Within the ESP sessions, the tools provided a common language and focus for analyzing and discussing teaching and learning; specifically, the level of cognitive demand, selection and implementation of cognitively challenging tasks. The tools enabled the ideas that emerged in the ESP sessions to "travel" into teachers' classrooms, to support teachers' selection and implementation of high-level tasks and to focus teachers' analysis of their own instructional practice. During the case stories, the tools served to frame conversations between teachers about their own attempts at implementing high-level tasks. According to the National Academy of Education (1999), "tools—including student assessments, curriculum and

professional-development materials…, and protocols for observing classrooms or professional meetings—are powerful carriers of theory and knowledge. Carefully designed tools that educators find useful in their practice can, then, become a powerful means of changing educational practice." The ESP workshop provided ESP teachers with tools that: 1) increased their awareness of the cognitive demands of mathematical tasks and of the influence of high-level tasks in supporting students' learning; 2) supported teachers' selection and implementation of cognitively challenging tasks in their own classrooms; 3) focused teachers' analysis and reflection on the implementation of high-level tasks in practice-based professional development materials and in their own practice; and 4) facilitated conversations between teachers about the implementation of high-level tasks in practice-based professional development materials and in their own classrooms. Through the consistent focus on the selection and implementation of high-level tasks, the tools provided to ESP teachers were useful in their practice and provided a powerful means of changing their practice.

Though effective, the ESP workshop could be improved in ways that would further enhance teachers' ability to select and implement cognitively challenging tasks in their own classrooms. An interesting finding is that none of the tasks that were rated as low-level for *Potential* in the student work collection or lesson observations increased to a high-level score for *Implementation*. A discussion during the sixth ESP session sheds light on this finding. Participants were commenting on the "Thinking through a Lesson" protocol (Hughes & Smith, 2004), and one participant (Cara) stated that the protocol was useful for high-level tasks but not for an "everyday lesson." This generated a discussion on how to make everyday instruction focus on meaning and understanding:

> Facilitator: "Does this suggest that a high-level task can't be an everyday lesson? So you have occasions where you do stuff like the EPT task [a pattern-generalization task] and you have days where you learn FOIL [the procedure for multiplying binomials]? Is that just the way it is, or is there a way to think about high-level tasks as being more integrated, more pervasive?
>
> Dave: I just thought you were going to ask the questions the opposite way; is there a way to make the day-to-day more high-level? …That's what I have been wrestling with all year in my algebra class (video transcript, Session 6, 5/07/05).

The discussion continues, lasting almost 14 minutes, with contributions from three other teachers and the following suggestion from the facilitator:

> One way to think about it is, is there a way to start a unit that you're working on in some way that can be higher-level so that you have some kind of conceptual underpinnings. Then when you do something that is more formulaic or procedurally-driven, at least you can always connect it back to something that has a conceptual foundation…. If you can connect that procedure to something that helps give it meaning, there is a greater chance that students will remember it and be able to use it in situations where it is appropriate (video transcript, Session 6, 5/07/05).

Interestingly, ten teachers referred to this discussion in the session evaluation, and three teachers referred to it in their post-workshop interview approximately one month later. In terms of the TAG, the pressing question concerns providing opportunities for students to make connections for tasks or procedures that are typically presented as procedures without connections or memorization. Note that, in the last session of the workshop, teachers were still wrestling with the idea of "procedures with connections;" which was the most frequently missed

category of task on the pre- and post-workshop task sort. Reflecting on this discussion and its significance to a majority of the teachers, adapting procedural tasks to have the potential to connect to meaning and understanding was not adequately addressed within the ESP workshop.

The following section will describe how the ESP workshop and the methodology utilized in this study builds upon prior professional development research and can inform future professional development for teachers of mathematics.

5.3. Contributions of this Investigation: Utilizing and Extending Prior Professional Development Research

This study contributes to a growing body of research on the design of effective professional development for teachers of mathematics and on the study of teachers' learning and instructional change following their participation in professional development experiences. To begin with, the results of this study lend further credence to the effectiveness of applying social-constructivism to teacher-learning, as utilized in several professional development studies (i.e., Farmer, et al., 2003; Simon, et al., 2000; Smith, 2000; Cobb, et al., 1991; Simon & Shifter, 1991). In this investigation, a social-constructivist approach to the design and facilitation of the ESP workshop appeared to be successful in supporting changes in ESP teachers' knowledge and beliefs about effective mathematics teaching and learning. Situating teachers' learning in practice-based professional development materials (Smith, 2001; Ball & Cohen, 1999) also contributed to teachers' opportunities for learning within the ESP workshop. Specifically, as suggested by prior studies (i.e., Shifter & Simon, 1992; Borasi, et al., 1999), engaging teachers in solving cognitively challenging tasks and prompting them to explicitly reflect on their own learning was a valuable tool for eliciting teacher-generated ideas about the selection and

implementation of high-level tasks throughout the ESP workshops. These ideas were prominent in teachers' post-workshop criteria for high-level tasks, in their evaluations of their own learning from the ESP sessions, and in their written reflections from the ESP sessions. As encapsulated in the comments of several ESP teachers, the use of narrative cases also appeared to foster teachers' examination of and reflection on issues of task implementation in their own classrooms. The finding that teachers who frequently commented on issues of implementation were also the teachers who exhibited the greatest degree of instructional change supports the contention that the analysis of narrative cases can influences teachers' own instructional practices as suggested by scores of theorists and researchers (e.g., Wallen & Williams, 2000; Barnett, 1998, 1991; Shulman, 1992; Sykes & Byrd, 1992; for a comprehensive review, see Merseth, 1996).

The current investigation strengthens the knowledge base of teachers' instructional change following their participation in professional development activities by describing changes in teachers' implementation of high-level tasks and by utilizing classroom artifacts and observations as the main data source. Teachers in several professional development studies increased the use of high-level tasks in their own classrooms (i.e., Swafford et al., 1997; Farmer, et al., 2003; Borasi, et al., 1999), as did ESP teachers. This study extends earlier research by analyzing student work and lesson observations to provide evidence that teachers also improved their ability to maintain the high-level cognitive demands during implementation. A distinguishing feature of this investigation is the utilization of a tool for analyzing classroom observations and collections of student work (i.e., the IQA Academic Rigor in Mathematics rubric) that provided descriptive information and served as a statistically sound instrument for collecting quantitative data on teachers' selection and implementation of cognitively challenging tasks. Hence, statistically significant increases in teachers' selection and implementation of high-

level tasks could be identified, and these changes could be described in ways that portrayed what the differences "looked like" in teachers' classrooms or in students' work.

This study also contributes to research on the use of student work as an indicator of classroom practice (Matsumura, et. al., 2002; Clare & Aschbacher, 2001). Recent research by the IQA team has shown that in mathematics, student work is a stable measure of instructional practice that is highly correlated with observed instruction (Matsumura, Slater, Junker, Peterson, Boston, Steele, & Resnick, 2006). As evidenced in the case studies of Randy and Nellie, features of students' work were very closely aligned with features of the lesson observations. This suggests that student work provides a proxy for lesson observations that is statistically and qualitatively consistent with observed instruction, and thus holds implications for the design of future research into teachers' instructional practices.

This investigation also drew on the methodology and frameworks generated by prior professional development research. CGI (Carpenter, et al., 1989) and QUASAR (Silver & Stein, 1996) provided models of professional development research that analyzed classroom artifacts and observations for evidence of change in teachers' instructional practices. CGI researchers designed professional development experiences based on a research framework, shared this framework with teachers, and then used the same framework as a tool for analyzing teachers' instructional practices. In this investigation, such alignment between the content and goals of the professional development activities, the objectives for teachers' learning and instructional change, and the instrument used to assessment teacher learning and instructional change provided a basis for connecting changes in teachers' knowledge and instructional practices to their experiences within the professional development sessions. Though not establishing causal links, the strong connections between changes in teachers' knowledge and instructional practices

and their experiences in the ESP workshop provide indications that learning occurred during the ESP workshop, and this learning may have influenced subsequent changes in teachers' classrooms.

The effectiveness of the professional development intervention and of the tools for studying teacher-learning and instructional change in this study can be attributed in large measure to the QUASAR project (Silver & Stein, 1996). Frameworks pioneered by QUASAR researchers provided the foundation for multiple aspects of this investigation:

1) the content and activities of the ESP professional development workshop (i.e., the TAG [Figure 2.1], the MTF [Figure 2.2], the cases created by Stein and colleagues [Stein, et al., 2000]; the factors and patterns identified by Stein, Grover & Henningsen [1996]);
2) the research on teachers' learning and instructional change (i.e., the task sort (Smith, et al., [2004]; the MTF; specifically, studying teachers instructional practices by comparing the level of the task vs. the level of implementation[3]); and
3) the structure and content of the data collection tool (i.e., the IQA Academic Rigor in Mathematics rubrics [Boston & Wolf, 2004, 2006]).

The effectiveness of using the QUASAR frameworks as the basis of professional development experiences for mathematics teachers is evident in the results of this study. While it is impossible to tease out which specific activities or aspects of the ESP professional development workshop had the greatest impact on teachers' learning and instructional change, the QUASAR frameworks served as the coherent thread that connected all of the individual

[3] Analyzing mathematics teachers' instructional practices by comparing the level of the task vs. the level of task implementation was also utilized in the recent TIMSS 1999 Video Study (USDE-NCES, 2003).

activities. These frameworks resonated with teachers' experiences in the classroom, provided a lens through which teachers came to view their own instructional practices (discussed later in this chapter), and provided a framework though which instructional change could be identified, quantified, and described (i.e., the IQA rubrics).

The closing section of this chapter will state conclusions that can be drawn from the study and suggest directions for continued research.

5.4. Conclusions and Directions for Future Research

The results of this investigation are important for several reasons. First and foremost, teachers in the study improved their instructional practices along dimensions of teaching that have been linked to increase students' opportunities for leaning. Following their participation in the professional development workshop, ESP teachers were selecting cognitively challenging tasks more frequently and were more frequently implementing these tasks in ways that engaged students in high-level thinking and reasoning. Teachers appeared to increase their knowledge of the ways in which high-level tasks and high-level implementation support students' learning, and subsequently selected more high-level tasks and improved their ability to maintain the cognitive demands during implementation. Future research will endeavor to directly establish the link between teachers' knowledge of the cognitive demands of mathematical tasks, teachers' use of high-level mathematical tasks, and student learning outcomes

Second, the effectiveness of the ESP professional development workshop merits further investigation. Will the ESP workshop 'travel' to other situations? Can the results be replicated with a larger group of teachers, with elementary teachers, or with other facilitators? Will

improvements to the content of the workshop further enhance changes in teachers' instructional practices? These questions pose a rigorous agenda for future research.

Third, this investigation used quantitative methods to analyze teacher learning and instructional change, with descriptive data provided to support and instantiate the differences identified by statistical tests. Self-reports and teacher reflections were used to support and illustrate changes identified by classroom artifacts, observations, and participation in the professional development as recorded on videotape. Future research endeavors would seek a larger sample size to enable more statistical tests, and would include multiple dimensions of the IQA Academic Rigor rubrics to provide the potential for the statistical and descriptive assessment of a greater variety of teachers' instructional practices.

In summary, this study has provided data on the effectiveness of a specifically focused professional development workshop in improving teachers' knowledge, selection, and implementation of cognitively challenging tasks. These instructional changes hold promise for improving students' learning of mathematics in ESP teachers' classrooms, and suggest that the ESP workshop can serve as one model of the type of professional development capable of improving teachers' instructional practices and students' learning more broadly. Future directions for this study include a follow-up assessment of the maintenance and continued growth of ESP teachers' selection and implementation of high-level instructional tasks. Furthermore, future research endeavors will seek to replicate the results of the study with other groups of teachers and to directly establish the link between professional development, teachers' knowledge and instructional practices, and students' achievement in mathematics.

APPENDIX 3.1

SUMMARY OF ESP PROFESSIONAL DEVELOPMENT ACTIVITIES

Comparing Martha's Carpeting Task and the Fencing Task

Teachers solve two tasks similar in mathematical content but with different levels of cognitive demand, "Martha's Carpeting" task and "The Fencing Task" (Stein, et al., 2000). Teachers then compare the similarities and differences of the two tasks and the opportunities each task provides for students' learning.

Task Sort

Teachers engage in analyzing a set of tasks that differ with respect to their cognitive demands and task features (e.g., require an explanation, utilize a diagram, provide tools such as calculators). Intended to cause teachers to focus on the different opportunities for learning provided by mathematical tasks with different levels of cognitive demand. In small groups, teachers classify tasks as high or low level and provide rationale for their classification. The whole group then co-constructs criteria for high-level and low-level tasks and discusses why the ability to make this distinction is important for teachers of mathematics.

Case Discussions

Case discussions begin by engaging teachers in solving the mathematical task featured in the lesson portrayed in the case. Teachers then read the case and engage in small- and large-

group discussions about aspects of teaching and learning of mathematics that occurred in the case and that are important to teaching and learning mathematics more broadly. Specifically, teachers solve the "Linking Fractions, Decimals, and Percents" task featured in "The Case of Ron Castleman and the "Multiplying Monomials and Binomials" task featured in "The Case of Monique Butler" (Stein, et al., 2000).

Case Stories

Case stories (Ackerman, Maslin-Ostrowski, & Christensen, 1996) are a structured format for teachers to share their teaching practice with a small group of colleagues and for colleagues to provide feedback. Teachers are asked to teach a lesson using a pedagogical 'tool' highlighted within the previous ESP session and are provided with specific prompts to reflect on the teaching and learning that occurred in the lesson. Teachers return to the next ESP session with their written reflections and any evidence or artifacts (i.e., student work, lesson plans, transcribed or paraphrased interactions, video- or audio-taped segments of the lesson) to tell the 'story' of the lesson. "I noticed" & "I wondered" format prepares mentor teachers for facilitating non-threatening instructional conferences with their student teachers.

"Extend Pattern of Tiles" Task and Student Work

Teachers solve the NAEP released item "Extend Pattern of Tiles" (EPT task) and analyze samples of student work. Teachers are asked to identify the student work samples that illustrate the greatest and least understanding of the main mathematical ideas in the task. As a whole-group, teachers then explicate the criteria for a response that would illustrate the highest level of understanding, for this specific task and for open-ended mathematical tasks in general.

Teachers also analyze student work from the EPT task and determine what questions they would ask to the student who produced each sample of work to assess and advance the students' understanding of the main mathematical goals of the task. As a whole group, teachers then look across the 'assess' and 'advance' questions created in the small groups to identify general characteristics of assess and advance questions and to discuss why each type of questions is important to students' learning of mathematics.

Based on the student work from the EPT task, teachers work in small groups to select and sequence the student responses that they would have presented during a whole-group discussion of the task. Each small group provides a rationale for their selection and sequence, and the whole group then generalizes "rules of thumb" for orchestrating whole-group discussions based on students' work.

Thinking Through a Lesson

Teachers are introduced to the "Thinking Through a Lesson Protocol" (TTLP) (Hughes & Smith, 2004) used for lesson-planning in the Math Methods courses at Pitt. Teachers then use the TTAL protocol to plan a lesson based on a high-level mathematical task. This activity serves as a culminating activity for the Analyzing Teaching and Learning Workshop (the protocol encapsulates the activities in which teachers engaged throughout the workshop) and as a transitional activity into the Leadership & Mentoring Workshop (the mentors will use the TTAL to structure instructional conferences with their student teachers).

APPENDIX 3.2

THE MIDDLE-SCHOOL TASK SORT

TASK A

Manipulatives/Tools Available: Calculator

Treena won a 7-day scholarship worth $1,000 to the Pro Shot Basketball Camp. Round-trip travel expenses to the camp are $335 by air or $125 by train. At the camp she must choose between a week of individual instruction at $60 per day or a week of group instruction at $40 per day. Treena's food and other expenses are fixed at $45 per day. If she does not plan to spend any money other than the scholarship, what are <u>all</u> choices of travel and instruction plans she could afford to make? Explain which option you think Treena should select and why.

TASK B

Manipulatives/Tools Available: Counters

This question requires you to show your work and explain your reasoning. You may use drawings, words, and numbers in your explanation. Your answer should be clear enough so that another person could read it and understand your thinking. It is important that you show *all* your work.

A pattern of dots is shown below. At each step, more dots are added to the pattern. The number of dots added at each step is more than the number added in the previous step. The pattern continues infinitely.

(1st step)	(2nd step)	(3rd step)
		● ● ● ●
	● ● ●	● ● ● ●
● ●	● ● ●	● ● ● ●
2 dots	6 dots	12 dots

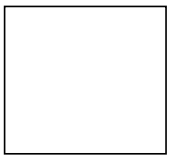
Marcy has to determine the number of dots in the 20th step, but she does not want to draw all 20 pictures and then count the dots.
Explain how she could do this *and* give the answer that Marcy should get for the number of dots.

TASK C

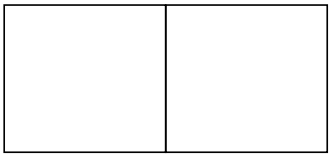
Manipulatives/Tools: Square Pattern Tiles

Using the side of a square pattern tile as a measure, find the perimeter (i.e., distance around) of each train in the pattern block figure shown below.

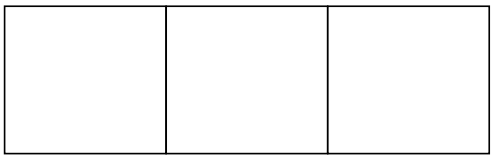
Train 1



Train 2



Train 3



TASK D

Manipulatives/Tools: None

Part A: After the first two games of the season, the best player on the girl's basketball team had made 12 out of 20 free throws. The best player on the boys' basketball team had made 14 out of 25 free throws. Which player had made the greater percent of free throws?

Part B: The "better" player had to sit out the third game due to an injury. How many baskets (out of an additional 10 free throw "tries") would the other player need to make in order take the lead in terms of greatest percentage of free throws?

TASK E

Manipulatives/Tools: Calculator

Divide using paper and pencil. Check your answer with a calculator and round the decimal to the nearest thousandth.

$$\frac{525}{1.3}$$

$$\frac{52.75}{7.25}$$

$$\frac{30.459}{.12}$$

TASK F

Manipulatives/Tools: None

Match the property name with the appropriate equation.

1. Commutative property of addition	a. $r(s+t) = rs + rt$
2. Commutative property of multiplication	b. $x \bullet 1/x = 1$
3. Associative property of addition	c. $-y + x = x + (-y)$
4. Associative property of multiplication	d. $a/b + -a/b = 0$
5. Identity property of addition	e. $y \bullet (zx) = (y\ z) \bullet x$
6. Identity property of multiplication	f. $1 \bullet (xy) = xy$
7. Inverse property of addition	g. $d \bullet 0 = 0$ and $0 \bullet d = 0$
8. Inverse property of multiplication	h. $x + (b + c) = (x + b) + c$
9. Distributive property	i. $y + o = y$
10. Property of zero for multiplication	j. $p \bullet q = q \bullet p$

TASK G

Manipulatives/Tools Available: Base Ten Blocks, grid paper

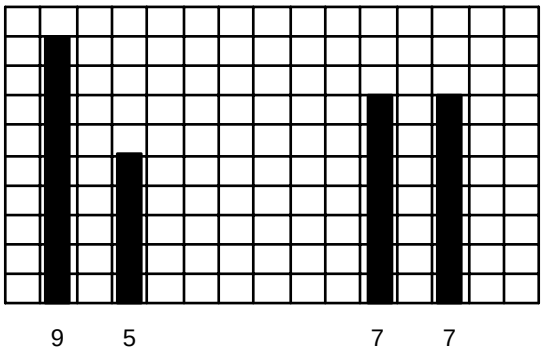
.08 .8 .080 .008000

Make three observations about the relative size of the above 4 numbers. Be sure to explain your observations as clearly as possible. Feel free to illustrate your observations if you feel it would help others understand them.

TASK H

Manipulatives/Tools: Grid Paper

The pairs of numbers in a - d below represent the heights of stacks of cubes to be leveled off. On grid paper, sketch the front views of columns of cubes with these heights before and after they are leveled off. Write a statement under the sketches that explains how your method of leveling off is related to finding the average of the two numbers.



By taking 2 blocks off the first stack and giving them to the second stack, I've made the two stacks the same. So the total # of cubes is now distributed into 2 columns of equal height. And that is what average means.

a) 14 and 8 b) 16 and 7 c) 7 and 12 d) 13 and 15

TASK I

Manipulatives/Tools: None

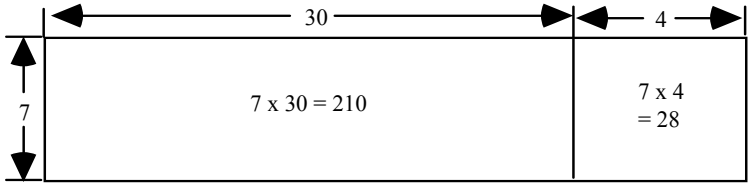
Write and solve a proportion for each.

17 is what percent of 68?
What is 15% of 60?
8 is 10% of what number?
24 is 25% of what number?
28 is what percent of 140?
What is 60% of 45?
36 is what percent of 90.
What is 80% of 120?
21 is 30% of what number?

TASK J

Manipulatives/Tools: None

One method of mentally computing 7 x 34 is illustrated in the diagram below:



Mentally compute these products. Then sketch a diagram that describes your methods for each.
a) 27 x 3

b) 325 x 4

TASK K

Manipulatives/Tools Available: Calculator with scientific functions

Penny's mother told her that several of her great-great-great-grandparents fought in the Civil War. Penny thought this was interesting and she wondered how many great-great-great grandparents that she actually had. When she found that number, she wondered how many generations back she'd have to go until she could count over 100 ancestral grandparents or 1000, or 10,000, or even 100,000. When she found out she was amazed and she was also pretty glad she had a calculator. How do you think Penny might have figured out all of this information? Explain and justify your method as clearly and completely as possible.

TASK L

Manipulatives/Tools: Base-10 Blocks

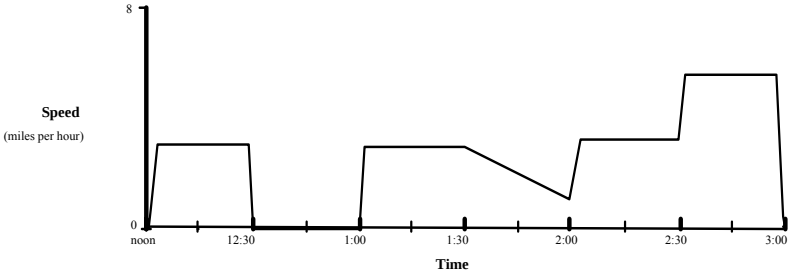
Using Base-10 blocks, show that 0.292 is less than 0.3.

TASK M

Manipulatives/Tools Available: None

Use the following information and the graph to write a story about Tony's walk:

At noon, Tony started walking to his grandmother's house. He arrived at her house at 3:00. The graph below shows Tony's spe in miles per hour throughout his walk.



Write a story about Tony's walk. In your story, describe what might have been doing at the different times.

TASK N

Manipulatives/Tools: None

The cost of a sweater at J. C. Penney's was $45.00. At the "Day and Night Sale" it was marked 30% off of the original price. What was the price of the sweater during the sale? Explain the process you used to find the sale price.

TASK O

Manipulatives/Tools: None

Give the fraction and percent for each decimal.

.20 = ____=____

.25 = ____=____

$.3\overline{3}$ = ____=____

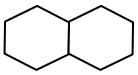
.50 = ____=____

$.6\overline{6}$ = ____=____

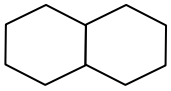
.75 = ____=____

TASK P

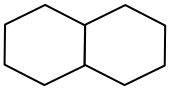
Manipulatives/Tools: Pattern Blocks

For problems 1-3, use  as the whole or unit.

1. Find 1/2 of 1/3. Use pattern blocks. Draw your answer.

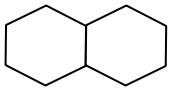


Show 1/3.

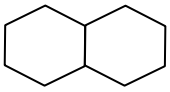


Show 1/2 of 1/3. 1/2 x 1/3 = ☐

2. Find 1/3 of 1/4. Use pattern blocks. Draw your answer.

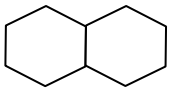


Show 1/4.

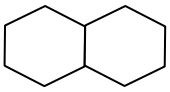


Show 1/3 of 1/4. 1/3 x 1/4 = ☐

3. Find 1/4 of 1/3. Use pattern blocks. Draw your answer.



Show 1/3.



Show 1/4 of 1/3. 1/4 x 1/3 = ☐

APPENDIX 3.3

PROTOCOL FOR TASK SORT

[Have tasks on individual cards. Give RED pen to write with.]

The 16 tasks on these cards are taken from middle school mathematics curricular materials. Working on your own, without talking to your neighbors, I'd like you to sort the tasks into two categories that we are calling high level and low level -- and we would like you to develop a list of criteria for high and low level tasks. I am interested in how you are deciding whether a task is H or L level. Notice on the back of each card, there is a place to indicate which category you have placed the task in (as well as a category of 'not sure') and a space to provide a brief rationale as to why you choose H-L or L-L (or unsure) for that particular task.

Once you have the tasks sorted, there are also cards for you to describe your criteria for including a task in the high-level category and for including a task in the low level category.

I will be around to answer any questions on an individual basis. Again, please work individually without consulting other members of your table. We will have a chance to discuss our categories this afternoon. After 20 minutes, I will check in to see where everyone is at.

When you are finished, (1) on the recording sheet, indicate which tasks

you placed in each category; (2) then place all of the cards in your envelope for safe-keeping until later.

[Collect RED pens]

APPENDIX 3.4

DIRECTIONS FOR TASK COLLECTION

- Identify 5 consecutive days of instruction (within the same chapter/unit) between now and (date). If you are being observed, the lesson observation needs to be included within the 5 days. Please do not include days in which the majority of students' time is spent taking a test, quiz, or other type of assessment.

- Please make copies of all of the mathematical tasks you use for any purpose during the 5 consecutive days of instruction. "Mathematical tasks" include any mathematical problems, exercises, examples, or individual or group work that students encounter from when the bell rings to begin the class period until the bell rings to end the class period.

- Please place the copies of the tasks in the file marked for the appropriate day. For each day, number the tasks according to their order in the day's lesson. On the log sheet provided in each day's folder, indicate the source of the task, approximately how much time was spent on the task and what purpose the task served in the lesson. For example, the task might have been used:
    - as a "warm-up" or "problem of the day"
    - to introduce the math ideas in the day's lesson
    - to develop the math ideas in the day's lesson
    - as independent or group work during class
    - as an assignment

- Please include a copy of a lesson plan for at least 1 of the 5 task-collection days. The lesson plan can consist of anything that you typically create or write down in preparation for a lesson. If you are being observed, include the lesson plan for the observed lesson.

APPENDIX 3.5

TASK LOG SHEET

Day ____ Teachers' Initials _____

TASK #	**SOURCE** of the task	**TIME SPENT** on the task	**PURPOSE** of the task in the lesson

APPENDIX 3.6

STUDENT WORK COVER SHEET

Task # _____ on Day _____

1. Indicate if this assignment is **typical** ☐. If not, please explain:

2. Describe any instructions or directions that were given to students:

3. How did you structure students' work on the task? [What did you do? What did students do?]:

4. How did you assess students' work on the task? [What did you expect to see in students' work on the task? What products/processes were students held accountable for?]:

APPENDIX 3.7

DIRECTIONS FOR STUDENT WORK COLLECTION

- Collect class-sets of student work (i.e., the written work from each student or group of students) for 3 of the tasks within the 5 consecutive days of instruction. The 3 class-sets of student work should be from different days. Please do not include students' tests or quizzes.

- Please make copies of the students' work with the students' names removed.

- Complete a Student Work Cover Sheet for each class-set of student work.

- From each class-set of student work, identify:
    - 2 samples of high-quality work (mark with the BLUE stickers provided)
    - 2 samples of medium-quality work (mark with the RED stickers)
    - 2 samples of low-quality work (mark with the GREEN stickers)

- Please place each set of student work and the Student Work Cover Sheet in the files marked for Student Work 1, Student Work 2, and Student Work 3.

APPENDIX 3.8

IQA ACADEMIC RIGOR: MATHEMATICS RUBRICS

RUBRIC 1: Potential of the Task

4	**The task has the potential to engage students in exploring and understanding the nature of mathematical concepts, procedures, and/or relationships, such as:** • Doing mathematics: using complex and non-algorithmic thinking (i.e., there is not a predictable, well-rehearsed approach or pathway explicitly suggested by the task, task instructions, or a worked-out example); OR • Procedures with connections: applying a broad general procedure that remains closely connected to mathematical concepts. **The task must explicitly prompt for evidence of students' reasoning and understanding.** For example, the task **MAY** require students to: • solve a genuine, challenging problem for which students' reasoning is evident in their work on the task; • develop an explanation for why formulas or procedures work; • identify patterns and form 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; • 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 **The task does not require student to engage in cognitively challenging work; the task is easy to solve.**
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.** **OR The task requires no mathematical activity.**

RUBRIC 2: IMPLEMENTATION OF THE TASK

4	Students engaged in exploring and understanding the nature of mathematical concepts, procedures, and/or relationships, such as: • Doing mathematics: using complex and non-algorithmic thinking (i.e., there is not a predictable, well-rehearsed approach or pathway explicitly suggested by the task, task instructions, or a worked-out example); OR • Procedures with connections: applying a broad general procedure that remains closely connected to mathematical concepts. There is explicit evidence of students' reasoning and understanding. For example, students may have: • solved a genuine, challenging problem for which students' reasoning is evident in their work on the task; • developed an explanation for why formulas or procedures work; • identified patterns and formed 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 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); • students identified patterns but did not make 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
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. S**tudents 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 **Student did not engage in cognitively challenging work; the task was easy to solve.**
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.** **OR Students did not engage in mathematical activity.**

APPENDIX 3.9

IQA LESSON CHECKLIST

Check each box that applies:

A ↑ The Lesson provided **opportunities** for students **to engage** with the **high-level demands** of the task:	**B** ↓ During the Lesson, the high-level **demands of the task** were **removed or reduced**:
▫ Students engaged with the task in a way that addressed the teacher's goals for high-level thinking and reasoning. ▫ Students communicated mathematically with peers. ▫ Students had appropriate prior knowledge to engage with the task. ▫ Teacher supported students to engage with the high-level demands of the task while maintaining the challenge of the task ▫ Students had opportunities to serve as the mathematical authority in the classroom. ▫ Teacher provided sufficient time to grapple with the demanding aspects of the task and for expanded thinking and reasoning. ▫ Teacher held students accountable for high-level products and processes. ▫ Teacher provided consistent presses for explanation and meaning. ▫ Teacher provided students with sufficient modeling of high-level performance on the task. ▫ Teacher provided encouragement for students to make conceptual connections. ▫ Students had access to resources that supported their engagement with the task. ▫ Other:	▫ The task expectations were not clear enough to promote students' engagement with the high-level demands of the task. ▫ The task was not complex enough to sustain student engagement in high-level thinking. ▫ The task was too complex to sustain student engagement in high-level thinking (i.e., students did not have the prior knowledge necessary to engage with the task at a high level). ▫ Classroom management problems interfered with students' opportunities to engage in high-level thinking. ▫ Teacher provided a set procedure for solving the task ▫ The focus shifted to procedural aspects of the task or on correctness of the answer rather than on meaning and understanding. ▫ Feedback, modeling, or examples were too directive or did not leave any complex thinking for the student. ▫ Students were not pressed or held accountable for high-level products and processes or for explanations and meaning. ▫ Students were not given enough time to deeply engage with the task or to complete the task to the extent that was expected. ▫ Students did not have access to resources necessary to engage with the task at a high level. ▫ Other:

APPENDIX 3.10

SCORING MATRIX FOR TASK SORT

Task Sort Item	Present in Participant's Written Response (Score = 1)	Not Present in Participant's Written Response (Score = 0)
Identifying the Level of Cognitive demand of the Task (High or Low)	Participant has selected correct level of cognitive demand	Participant has selected incorrect level of cognitive demand or "Not Sure"
Provide Rationale for the selected level of cognitive demand of the task	Participant identifies elements of the task that are consistent with descriptors in the TAG or synonymous.	Participant identifies elements of the task that do not reflect the task's potential to provide opportunities for high-level thinking and reasoning (i.e., surface level features or characteristics in conflict with the TAG)
List Criteria for High Level Tasks	Participant identifies the category of doing mathematics or the descriptors of doing mathematics tasks in the TAG.	Participant does not identify the category of doing mathematics or the descriptors of doing mathematics tasks in the TAG.
	Participant identifies the category of procedures with connections or the descriptors of procedures with connections tasks in the TAG.	Participant does not identify the category of procedures with connections or the descriptors of procedures with connections tasks in the TAG.
	Participant identifies surface-level features consistent with high-level task demands.	Participant identifies surface-level features inconsistent with high-level task demands.
List Criteria for Low Level Tasks	Participant identifies the category of procedures without connections or descriptors of procedures without connections tasks in the TAG.	Participant does not identify the category of procedures without connections or the descriptors of procedures without connections tasks in the TAG.
	Participant identifies the category of memorization, or the descriptors of memorization tasks in the TAG.	Participant does not identify the category of memorization or the descriptors of memorization tasks in the TAG
	Participant identifies surface- level features consistent with low-level task demands.	Participant identifies surface-level features inconsistent with low-level task demands.

APPENDIX 4.1

ATTRITION

Teachers providing data in each data collection

	Number of teachers submitting tasks	Identification of teachers not submitting tasks	Number of teachers submitting student work	Identification of teachers not submitting student work
Fall	18		16	A, B
Winter	16	C, D	15	A, C, D
Spring	14	C, E, F, G	13	A, C, E, F, G

Eighteen teachers provided a data collection packet in the Fall, though only 16 of these teachers submitted student work. The two teachers who did not submit student work (Teachers A and B) had not received permission from their school district to collect student work for the study at that point in time. Teacher B received permission and submitted student work in the Winter and Spring; Teacher A did not.

Sixteen teachers submitted task collections in the Winter data collection. For the two teachers did not submit tasks in the Winter, Teacher C had a Fall task mean of 3.2 (all 5 tasks were high-level) and Teacher D had a Fall task mean of 2.2. (1 of 5 tasks was high-level). Teachers C and D also did not submit student work in the Winter. In the Spring, Teacher D

submitted tasks and student work (task mean = 3.0; 4 of 5 tasks were high-level), though Teacher C did not.

Fourteen teachers submitted task collections in the Spring and four did not. In addition to Teacher A, three other teachers also did not submit task collections in the Spring (Teachers E, F, and G). The Winter task means for these three teachers were 3.7, 3.1, and 3.1, respectively. Hence, the increase in task means in over time was not due to the attrition of low-scoring teachers.

Attrition does not appear to be the result of a lack of "buy-in" to the professional development workshop, as the four teachers who did not submit task collections in the Spring were among the most active participants in the ESP workshop and frequently made verbal contributions. Note that the same teachers who did not submit task collections in the Winter and Spring also did not submit student work. Feedback from these teachers indicated that their teaching workload or other responsibilities prohibited them from submitting a data collection packet.

BIBLIOGRAPHY

Ackerman, R., Maslin-Ostrowski, P., & Christensen, C. (1996). Case stories: Telling tales about school. *Educational Leadership, 53*(6), 21-23.

Arbaugh, F. (2000). *Time on tasks: Influences of a study group on secondary mathematics teachers' knowledge, thinking, and teaching.* Unpublished doctoral dissertation, (Indiana University, Bloomington). UMI Dissertation Services, #9993602.

Arbaugh, F. & Brown, C.A. (2002). *Influences of the mathematical tasks framework on high school mathematics teachers' knowledge, thinking, and teaching.* Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA).

Ball, D.L., & Cohen, D.K. (1999). Developing practice, developing practitioners: Towards a practice-based theory of professional education. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 3-32). San Francisco: Jossey-Bass.

Ball, D.L., Lubenski S.T., & Mewborn, D.S. (2001). Research on teaching mathematics: The unresolved problem of teachers' mathematical knowledge. In V. Richardson (Ed.), *Handbook of research on teaching* (4th edition, pp. 433-456). Washington DC: American Educational Research Association.

Barnett, C. (1998). Mathematics teaching cases as a catalyst for informed strategic inquiry. *Teaching and Teacher Education, 14*(2), 81-93.

Barnett, C. (1991). Building a case-based curriculum to enhance the pedagogical content knowledge of mathematics teachers. *Journal of Teacher Education, 42*(4), 263-272.

Ben-Chaim, D., Fey, J. T., Fitzgerald, W. M., Benedetto, C., Miller, J. (1998). Proportional reasoning among 7th grade students with different curricular experiences. *Educational Studies in Mathematics, 36*(3), 247-273.

Brown, J. S., & VanLehn, K. (1982). Towards a generative theory of "bugs." In T. P. Carpenter, J. M. Moser, & T. A. Romberg (Eds.), *Addition and subtraction: A cognitive perspective* (pp. 117-135). Hillsdale, NJ: Erlbaum.

Borasi, R., Fonzi, J., Smith, C., Rose, B.J. (1999). Beginning the process of rethinking mathematics instruction: A professional development program. *Journal of Mathematics Teacher Education, 2*, 49-78.

Borko, H. & Putnam, R.T. (1995). Expanding a teacher's knowledge base: A cognitive psychological perspective on professional development. In T.R. Guskey & M. Huberman (Eds.), P*rofessional development in education: New paradigms and practices*. Mahwah, NJ: Lawrence Erlbaum.

Boston. M. D., & Wolf, M. K. (2006). *Assessing academic rigor in mathematics instruction: The development of Instructional Quality Assessment Toolkit*. Report submitted to the National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Boston. M.D., & Wolf, M.K. (2004). *Using the IQA toolkit to assess academic rigor in mathematics lessons and assignments*. Paper presented at the 2004 Annual Meeting of the American Educational Research Association (San Diego, CA).

Boston, M. (2003). *Taking advantage of openings in the curriculum to promote teachers' learning from cases.* Paper presented at the 2003 Annual Meeting of the American Educational Research Association (Chicago, IL).

Carpenter, T.P., Fennema, E., Peterson, P.L., Chiang, C., & Loef, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: An experimental study. *American Educational Research Journal, 26*(4), 499-531.

Clare, L. (2000). *Using teachers' assignments as an indicator of classroom practice* (CSE Tech. Rep. No. 532). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Clare, L., & Aschbacher, P. R. (2001). Exploring the technical quality of using assignments and student work as indicators of classroom practice. *Educational Assessment, 7*(1), 39-59.

Clarke, D.M. (1997). The changing role of the mathematics teacher. *Journal for Research in Mathematics Education, 28*(3), 278-308.

Cobb, P., Yackel, E., & Wood, T. (1991). Curriculum and teacher development: Psychological and anthropological perspectives. In E. Fennema, T. P. Carpenter, & S. J. Lamon (Eds.), *Integrating research on teaching and learning mathematics*. Albany, NY: State University of New York Press.

Cooney, T. J. (1999). Conceptualizing teachers' ways of knowing. *Educational Studies in Mathematics, 38*(1), 163-187.

Denzin, N. K., & Lincoln, Y. S. (2000). Introduction: The discipline and practice of qualitative research. In N. S. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research*, (pp. 1-28). Thousand oaks, CA: Sage Publications.

Doyle, W. (1988). Work in mathematics classes: The context of students' thinking during instruction. *Educational Psychologist, 23*(2), 167-180.

Doyle, W. (1983). Academic work. *Review of Educational Research, 53*(2), 159-199.

Farmer, J. D., Gerretston, H., & Lassak, M. (2003). What teachers take from professional development: Cases and implications. *Journal of Mathematics Teacher Education, 6*, 331-360.

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

Fennema, E. & Franke, M. L. (1992). Teachers' knowledge and its impact. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 147-164). New York: Macmillan.

Forman, E., Larreamendy-Joerns, J., Stein, M. K., & Brown, C. A. (1998). 'You're going to want to find out which and prove it': Collective argumentation in mathematics classrooms. *Learning and Instruction, 8*, 527–48.

Franke, M., Carpenter, T., 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.

Fuson, K. C., Carroll, W. M., and Druek, J. V. (2000). Achievement results for second and third graders using standards-based curriculum Everyday Mathematics. *Journal for Research in Mathematics Education, 31*(3), 277-295.

Greeno, J. G. (1991). Number sense as situated knowing in a conceptual domain. *Journal of Research in Mathematics Education, 25*(1), 94-99.

Grover, B. (1989). *Development and preliminary evaluation of a scoring system for a semi-structured interview for teacher assessment.* Unpublished doctoral dissertation, (University of Pittsburgh). UMI Dissertation Services, #9019383.

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.

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. 65-100). New York: Macmillian.

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

Hiebert, J., & Wearne, D. (1993). Instructional tasks, classroom discourse, and students' learning in second-grade arithmetic. *American Educational Research Journal, 30*(2), 393-425.

Hill, H. C. & Ball, D. L. (2004). Learning mathematics for teaching: Results from California's mathematics professional development institutes. *Journal for Research in Mathematics Education, 35*(5), 330-351.

Hill, H. C., Rowan, B. & Ball, D. L. (2005). Effects of teachers' mathematical knowledge on student achievement. *American Educational Research Journal, 42*(2), 371-406.

Hughes, E. K., & Smith, M. S. (2004). *Thinking through a lesson: Lesson planning as evidence of and a vehicle for teacher learning.* Poster presented at the Annual Conference of the American Educational Research Association (San Diego, CA).

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 function strand. *Journal for Research in Mathematics Education, 31*(3), 328-361.

Janesick, V. J. (2000). Choreography of qualitative research design: Minuets, improvisations, and crystallization. In N.S. Denzin & Y.S. Lincoln (Eds.), *Handbook of qualitative research*, (pp. 379-399). Thousand Oaks, CA: Sage Publications.

Junker, B., Matsumura, L. C., Crosson, A., Wolf, M. K., Levison, A., Weisberg, Y., & Resnick, L. (2004). *Overview of the Instructional Quality Assessment*. Paper presented at the annual meeting of the American Educational Research Association (San Diego, CA).

Lai, J. C. (1995). Adragogy of the oppressed: Emancipatory education for Christian adults. ERIC Clearinghouse (#ED396104).

Leinhardt, G. (1987). The development of an expert explanation: An analysis of a sequence of subtraction lessons. *Cognition and Instruction*, *4*(4), 225-282.

Leinhardt, G. & Smith, D. A. (1985). Expertise in mathematics instruction: subject matter knowledge. *Journal of Educational Psychology, 77*(3), 247-71.

Lesh, R. A., Post, T. R., & Behr, M. J. (1986). Representations and translations among representations in mathematics learning and problem solving. In C. Janvier (Ed.), *Problems of representation in the learning of mathematics* (pp. 153-176). Mahwah, NJ: Lawrence Erlbaum.

Lloyd, G. M. (1999). Two teachers' conceptions of a reform-oriented curriculum: implications for mathematics teacher development. *Journal of Mathematics Teacher Education, 2*(3), 227-52.

Lloyd, G. M., & Wilson, M. L. (1998). Supporting innovation: The impact of a teacher's conception of functions on his implementation of a reform curriculum. *Journal for Research in Mathematics Education, 29*(3), 248-274.

Lotan, R. A. (2003). Group-worthy tasks. *Educational Leadership, 60*(6), 72-75.

Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Marx, R. W., & Walsh. J. (1988). Learning from academic tasks. *Elementary School Journal, 88*(3), 207-219.

Matsumura, L. C. (2003). *Teachers' assignments and student work: Opening a window on classroom practice* (CSE Tech. Rep. No. 602). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Matsumura, L. C., Garnier, H., Pascal, J., & Valdes, R. (2002). Measuring instructional quality in accountability systems: Classroom assignments and student achievement. *Educational Assessment, 8*(3), 207-229.

Matsumura, L. C., Slater, S. C., Junker, B., Peterson, M., Boston, M., Steele, M., & Resnick, L., (2006). *Reading comprehension and mathematics instruction in urban middle schools: a pilot study of the instructional quality assessment toolkit*. Report submitted to the National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Matz, M. (1980). Towards a computational theory of algebraic competence. *Journal of Mathematical Behavior, 3*(1), 93-166.

Merseth, K. (1996) Cases and case methods in teacher education. In J. Sikula (Ed.), *Handbook of research on teacher education* (2nd edition, pp. 722-744). New York: MacMillan

National Academy of Education, (1999). *Recommendations regarding research priorities: An advisory report to the National Educational Research Policy and Priorities B*oard. National Academy of Education. Available at http://www.naeducation.org/1NAEd_Publications.

National Council of Educational Statistics, (2004). *The nation's report card: Mathematics highlights 2003*. http://www.nces.ed.gov/nationsreportcard/pdf/main2003/2004451.pdf

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

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: NCTM.

National Council of Teachers of Mathematics. (1991). *Professional standards for teaching mathematics*. Reston, VA: NCTM.

O'Connor, M.C., & Michaels, S. (1996). Shifting participant frameworks: Orchestrating thinking practices in group discussions. In D. Ghicks (Ed.), *Discourse, learning, and schooling* (pp. 63-103). New York: Cambridge University Press.

Pennsylvania Department of Education (2004). *2003 State report card*. http://www.pde.state.pa.us/pas/lib/pas/correctachievement012604.pdf

Perfetti, C.A. & Marron, M.A. (1998). Learning to read: Literacy acquisition by children and adults. In D.A. Wagner (Ed.), *Advances in adult literacy research and development* (pp. 1-43). Melbourne, Australia: Hampton Press.

Post, T., Harel, G., Behr, J., & Lesh, R. (1991). Intermediate teachers' knowledge of rational number concepts. In E. Fennema, T. P. Carpenter, & S. J. Lamon (Eds.), *Integrating research on teaching and learning mathematics* (pp. 177-198). Madison, WI: Center for Education Research.

Putnam, R.T, Lampert, M., & Peterson, P.L. (1990). Alternative perspectives on knowing mathematics in elementary schools. In C.B. Cazden, (Ed.), *Review of research in education* (Vol. 16, pp. 57-150). Washington, D.C.: AERA.

Remillard, J. (1999). Curriculum materials in mathematics education reform: A framework for examining teachers' curriculum development. *Curriculum Inquiry, 29*(3), 315-342.

Remillard, J. T., & Bryans, M. B. (2004). Teachers' orientation toward mathematics curriculum materials: Implications for teacher learning. *Journal for Research in Mathematics Education, 35*(5), 352-388.

Remillard, J.T., & Geist, P.K. (2002). Supporting teachers' professional learning by navigating openings in the curriculum. *Journal of Mathematics Teacher Education, 5*(1), 7-34.

Reys, R., Reys, B., Lapan, R., & Holliday, G. (2003). Assessing the impact of Standards-based middle grades mathematics curriculum materials on student achievement. *Journal for Research in Mathematics Education, 34*(1), 74-95.

Ridgeway, J.E, Zawojewski, J.S., Hoover, M.N., & Lambdin, D.V. (2003). Student attainment I the Connected Mathematics curriculum. In S.L. Senk & D.R. Thompson (Eds.), *Standards-based mathematics curricula: What are they? What do students learn?* (pp. 193-224). Mahwah, NJ: Lawrence Erlbaum.

Riordan, J. E., & Noyce, P. E. (2001). The impact of two standards-based mathematics curricula on student achievement in Massachusetts. *Journal for Research in Mathematics Education, 32*(4), 368-398.

Rogers, A. (2003). What's the difference? *Adults Learning, 15*(2), 15-17.

Romanagno, L. (1994). *Wrestling with change: The dilemmas of teaching real mathematics*. Portsmouth, NH: Heinemann.

Romberg, T.A., & Carpenter, T.P. (1986). Research on teaching and learning mathematics: Two disciplines of scientific inquiry. In M. C. Wintrock, (Ed.), *Handbook of research on teaching* (3rd edition, pp. 850-873). American Educational Research Association.

Schifter, D.A., & Simon, M. (1992). Assessing teachers' development of a constructivist view of mathematics learning. *Teacher and Teacher Education, 8*, 187-197.

Schmidt, W. (1996). *The Third International Math and Science Study (TIMSS*). Speech presented at the conference "Using the NCES to Guide the Evaluation, Selection, and Adaptation of Instructional Materials." Washington DC. Available at http://teech.terc.edu/lectures/lectures/schmidt_nrc96.htm

Schoen, H. L., Fey, J. T., Hirsch, C. R., & Coxford, A. F. (1999). Issues and opinions in the math wars. *Phi Delta Kappan, 80*(6), 444-53.

Schoenfeld, A. H. (2002). Making mathematics work for all children: Issues of standards, testing, and equity. *Educational Researcher, 31*(1), 13-25.

Schonfeld, A. (1998). Toward a theory of teaching-in-context. *Issues in Education, 4*(1), 1-95.

Sherin, M. G. (2002). When teaching becomes learning. *Cognition and Instruction, 20*(2), 119-150.

Shulman, L. S. (1992) Toward a pedagogy of cases. In J. Shulman (Ed.), *Case methods in teacher education* (pp. 1-30). New York: Teachers College Press.

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

Simon, M.A. & Shifter, D. (1991). Towards a constructivist perspective: An intervention study of mathematics teacher development. *Educational Studies in Mathematics, 22*(4), 309-331.

Simon, M.A., Tzur, R., Heinz, K., Kinzel, M., & Smith, M.S. (2000). Characterizing a perspective underlying the practices of mathematics teachers in transition. *Journal for Research in Mathematics Education, 31*(5), 579-601.

Smith, M.S. (2003). Developing teacher leaders in mathematics. *Math & Science Collaborative Journal, 9*, 39.

Smith, M.S. (2001). *Practice-based professional development for teachers of mathematics*. Reston, VA: NCTM.

Smith, M.S. (2000). Balancing the old and new: An experienced middle school teacher's learning in the context of mathematics instructional reform. *The Elementary School Journal, 100*(4), 351-375.

Smith, M.S. (1995). *A road to change*. Unpublished doctoral dissertation (University of Pittsburgh). UMI Dissertation Services, 9614231.

Smith, M. S., Boston, M., Steele, M. (2006). *Focusing on challenging mathematical tasks: A strategy for improving teaching and learning.* Paper presented at the 2006 Annual Meeting of the Association of Mathematics Teacher Educators (Tampa, FL).

Smith, M. S., & Stein, M. K. (2002). *ASTEROID: A study of teacher education: research on instructional design*. Grant proposal to the National Science Foundation.

Smith, M.S., & Stein, M.K. (1998). Selecting and creating mathematical tasks: From research to practice. *Mathematics Teaching in the Middle School, 3*(5), 344-350.

Smith, M. S., Stein, M. K., Arbaugh, F., Brown, C., & Mossgrove, J. (2004). Characterizing the Cognitive Demands of Mathematical Tasks: A Task Sorting Activity. In G. Bright and R. Rubenstein, (Eds.), *Professional development guidebook for perspectives on teaching of*

mathematics: Companion to the sixty-sixth yearbook (pp. 45-72). Reston, VA: National Council of Teachers of Mathematics.

Stein, M.K., Baxter, J. & Leinhardt G. (1990). Subject matter knowledge and elementary instruction: a case from functions and graphing. *American Educational Research Journal, 27*(4), 639-663.

Stein, M. K., Grover, B., & 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.

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., Smith, M. S., & Silver, E. A. (1999). The development of professional developers: Learning to assist teachers in new settings in new ways. *Harvard Educational Review, 69*(3), 237-269.

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.

Swafford, J. O., Jones, G. A., & Thornton, C. A. (1997). Increased knowledge in geometry and instructional practice. *Journal for Research in Mathematics Education, 28*(4), 467-483.

Sykes, G. & Bird, T. (1992). Teacher education and the case idea. In G. Grant (Ed.), *Review of research in education* (Volume 18, pp. 457-521). Washington DC: American Educational Research Association.

Thompson, D. R., & Senk, S. L. (2001). The effects of curriculum on achievement in second-year algebra: The example of the University of Chicago School Mathematics Project. *Journal for Research in Mathematics Education, 32*(1), 58-84.

Thompson, C.L. & Zeuli, J.S. (1999). The frame and the tapestry: Standards-based reform and professional development. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 341-375). San Francisco: Jossey-Bass.

Thompson, A. (1992). Teachers' beliefs and conceptions: A synthesis of the research. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 127-146). New York: Macmillan.

USDE (1999). *Exemplary and promising mathematics programs: Expert panel report*. www.enc.org/professional/federalresources/exemplary/promising/

USDE (2000). *Before it's too late. A report to the Nation from the National Commission on Mathematics and Science Teaching for the 21st century*. Washington, DC: Author.

USDE, NCES (2003). *Teaching mathematics in seven countries: Results from the TIMSS 1999 video study.* NCES (2003-013). Washington, DC: Author.

Van de Walle, J. (2004). *Elementary and middle school mathematics: teaching developmentally* (4th edition). New York: Longman.

Voigt, J. (1994). Negotiation of mathematical meaning and learning mathematics. *Educational Studies in Mathematics, 26*(2), 275-298.

Wallen, S.B., & Williams, S.R. (2000). Validating classroom issues: Case methods in support of teacher change. *Journal of Mathematics Teacher Education, 3,* 3-26.

Weiss, I.R. & Pasley, J.P. (2004). What is high quality instruction? *Educational Leadership, 61*(5), 24-28.

Weiss, I.R.., Pasley, J.D., Smith, P.S., Banilower, E.R., & Heck, D.J. (2003). *Looking inside the classroom: A study of K–12 mathematics and science education in the United States (Highlights report).* Chapel Hill, NC: Horizon Research, Inc. Available at www.horizon-research.com.

Wood, T. (1995). An emerging practice of teaching. In P. Cobb & H. Bauresfeld (Eds.), *The emergence of mathematical meaning: Interaction in classroom cultures* (pp. 203-227). Hillsdale, NJ: Lawrence Erlbaum Associates.