J Math Teacher Educ (2013) 16:7–31
DOI 10.1007/s10857-012-9211-6

Connecting changes in secondary mathematics teachers' knowledge to their experiences in a professional development workshop

Melissa D. Boston

Published online: 10 March 2012

Abstract This investigation describes secondary mathematics teachers' learning and instructional change following their participation in a professional development workshop, the Enhancing Secondary Mathematics Teacher Preparation Project (ESP) (2004–2005), specifically focused on the selection and implementation of cognitively challenging mathematical tasks. Data consist of a pre/post-assessment of teachers' knowledge of the cognitive demands of mathematical tasks and videotaped discussions and written artifacts from the professional development sessions. A mixed methods approach was used to identify connections between teachers' learning and their experiences in the ESP workshop. Results indicate that ESP teachers developed new ideas about the influence of mathematical tasks on students' learning. Increases in teachers' knowledge of the cognitive demands of mathematical tasks were closely linked to ideas represented in frameworks and discussions from the ESP workshop and to teachers' experiences in solving challenging mathematical tasks as learners.

Keywords Professional development · Cognitive demands · Mathematical tasks · Teachers' learning · Instructional change

Introduction

Two decades ago, the National Council of Teacher of Mathematics (NCTM) unveiled standards for the teaching and learning of mathematics, proclaiming the importance of mathematical thinking, reasoning, and understanding in the lives and futures of students in American schools and portraying a vision of the type of mathematics teaching necessary to attain this goal (NCTM 1989, 1991). In this vision, teachers serve as facilitators of students' learning by providing opportunities for students to engage with rich mathematical tasks, develop connections between mathematical ideas and between different representations of mathematical ideas, and collaboratively construct and communicate their

M. D. Boston (✉)
Duquesne University, G-14 Canevin Hall, 600 Forbes Avenue, Pittsburgh, PA 15282, USA
e-mail: bostonm@duq.edu

mathematical thinking. However, large-scale national studies and comprehensive reviews of research provide evidence that this type of mathematics instruction has yet to be realized in the majority of US classrooms (Fullan 2009; Stein et al. 2007). A national assessment of mathematics teaching conducted by Horizon Research indicated that only 15 % of the 300 observed mathematics lessons provided students with opportunities to make connections, explore mathematics ideas, and develop mathematical understanding (Weiss et al. 2003). Similarly, results from the 1999 TIMSS video study identified several disheartening features of mathematics instruction in US classrooms: lack of coherence in mathematical ideas, low cognitive demands in 83 % of the mathematical tasks presented to students, and virtually no opportunities for students to make mathematical connections through a lesson (Stigler and Hiebert 2004). Less than 1 % of students' mathematical experiences involved "constructing relationships, …or engaging in mathematical reasoning such as conjecturing, generalizing, and verifying" (p. 98), and over half of instructional time was spent reviewing previously learned concepts or procedures in ways that did not advance the mathematical ideas.

In *Before It's Too Late,* the National Commission on Mathematics and Science Teaching for the twenty-first Century (USDE 2000) acknowledged the persistent need for improved mathematical learning in American schools 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)." Better teaching (i.e., moving teachers' practice toward the ideals of NCTM) will require teachers to engage in learning experiences that transform their knowledge and understanding of how mathematics is best taught and learned (Thompson and Zeuli 1999).

This article explores one promising route for improving mathematics teaching toward the ideals of NCTM. The *Enhancing Secondary Mathematics Teacher Preparation* (ESP) Project[1] provided secondary mathematics teachers with professional learning experiences focused on the selection and implementation of cognitively challenging mathematical tasks. Analyses of teachers' instructional practices before, during, and after their participation in ESP, based on classroom observations and collections of instructional tasks and students' work, indicated that teachers in the project significantly increased the use of cognitively challenging instructional tasks in their classrooms and their ability to implement these tasks in ways that maintained students' opportunities for thinking and reasoning (Boston 2006; Boston and Smith 2009). What remains unexplored, however, is the connection between teachers' experiences in the ESP Project, changes in teachers' knowledge, and the observed changes in teachers' instructional practices. Hence, this investigation examines the following questions:

1. In what ways did teachers' knowledge of the cognitive demands of mathematical tasks change following their participation in the ESP professional development workshop?
2. What is the relationship between changes in teachers' knowledge of the cognitive demands of mathematical tasks and their learning experiences in the ESP professional development workshop?

In this article, I argue for the importance of enhancing teachers' knowledge of the cognitive demands of mathematical tasks. I then describe the design of the ESP professional development workshop, present the research methods and results, and in discussing the results, pose a hypothesis connecting changes in teachers' knowledge to the changes in

[1] Principal investigators on the ESP Project were Margaret Smith, Ellen Ansell, Beverly Michaels, and Paul Gartside (University of Pittsburgh). The author served as ESP Project Director.

their instructional practices. I conclude the article by situating the importance of this work more broadly.

Background

The *Enhancing Secondary Mathematics Teacher Preparation* (ESP) project, funded by the National Science Foundation (NSF), created a series of professional development experiences to build teachers' capacity to engage 7–12th grade students in cognitively challenging mathematical activities. Improving instruction in ESP teachers' classrooms was intended to enhance the learning opportunities of their middle and high school students *and* of the prospective teachers who would be assigned as student teachers in their classrooms. To accomplish this goal, ESP teachers participated in three professional learning experiences over a two-year period, designed to: (1) support improvements in the teachers' own instructional practices (*Improving Practice*; 6 days in year 1); (2) develop the ability to mentor prospective and beginning teachers (*Becoming a Mentor*; 1 week at the end of year 1); and (3) promote a shared vision of effective mathematics teaching between teachers and the prospective teachers assigned to their classrooms (*Developing a Shared Vision*; five half-days in year 2).

In previously reported studies about the ESP Project, Smith and I identified changes in teachers' instructional practices following their participation in the year 1 "Improving Practice" workshop (Boston 2006; Boston and Smith 2009, 2011). At three points throughout the school year concurrent with this workshop, teachers submitted five consecutive days of instructional tasks, three class sets of students' written work (from within the five-day period), and one lesson observation (also within the five-day period). In the final data collection, teachers utilized significantly more instructional tasks with high-level cognitive demands and demonstrated an enhanced ability to maintain students' high-level thinking and reasoning. We hypothesized that the observed changes in teachers' instructional practices may have been catalyzed by teachers' new insights into the value of cognitively demanding instructional tasks in supporting students' learning of mathematics.

In this investigation, I explore teachers' learning and experiences in the year 1 "Improving Practice" workshop to identify what may have catalyzed the observed changes in teachers' instructional practices. Desimone (2009) proposes a "core conceptual framework for studying the effects of professional development on teachers and students" (p. 185). In this framework, Desimone suggests that professional development experiences generate change in teachers' knowledge, skills, attitudes, and beliefs, which then generate change in teachers' instructional practices and subsequently result in greater student learning. Specific to the ESP workshop, teachers' instructional change is envisioned to follow the conceptual model presented in Fig. 1. Professional development experiences focused on cognitively challenging mathematical tasks would enhance teachers' knowledge of task demands and of the impact of such tasks on students' learning. This new knowledge would then impact teachers' instructional practices, specifically their selection and implementation of cognitively challenging tasks, and subsequently serve to increase students' learning. Based on the observed changes in ESP teachers' instructional practices identified in previous work (e.g., Boston 2006, Boston and Smith 2009), this investigation aims to "trace" the changes in teachers' instructional practices back to teachers' learning and experiences in the ESP professional development workshop. To begin, I describe the importance of focusing teachers' learning on the cognitive demands of instructional tasks.

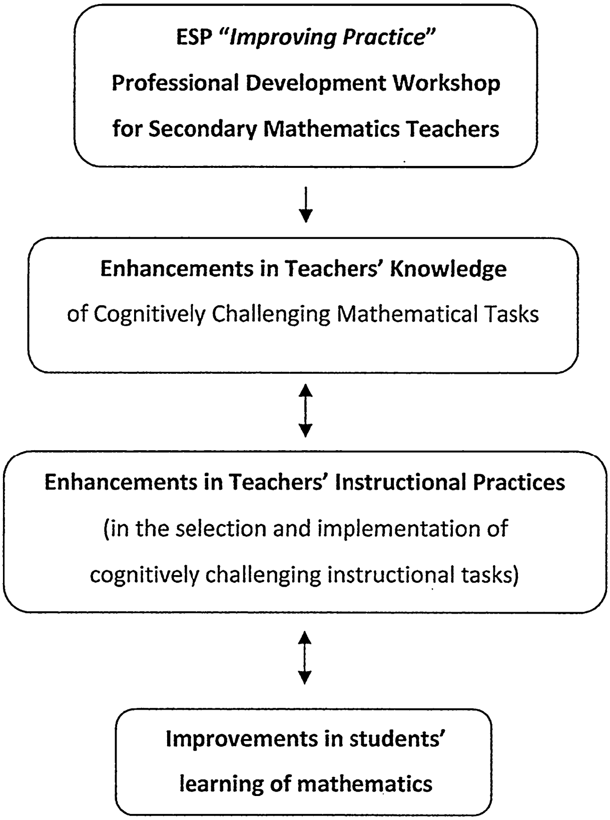

Fig. 1 Framework for connecting professional development to changes in teachers' knowledge and practice in the Enhancing Secondary Mathematics Teachers' Instructional Practices (ESP) Project (Boston and Smith 2008)

Increasing teachers' knowledge of the cognitive demands of instructional tasks

To impact students' learning, transformative learning experiences for teachers (i.e., experiences capable of changing teachers' underlying beliefs and conceptions of effective instruction [Thompson and Zeuli 1999]) should develop aspects of pedagogy empirically associated with increased mathematical understanding and achievement. While assessing student achievement in ESP teachers' classrooms was beyond the scope of this project,[2] consistent results across more than a decade of classroom research indicate that the *nature of instructional tasks* and *the way tasks are implemented during instruction* significantly influence students' opportunities to learn mathematics (e.g., Stein et al. 2007; Stigler and Hiebert 2004; Tarr et al. 2008). A "mathematical task" is defined as a mathematical problem or set of problems that address a specific mathematical idea (Stein et al. 2000). Mathematical tasks are situated "in the interactions of teaching and learning" (Stein et al. 2000, p. 25) as teachers select the tasks with which students engage during mathematics

[2] Teachers in the project were from different schools and school districts, teaching different grade levels and mathematics courses. Student achievement data and value-added data were not available to the ESP research team.

instruction, and tasks structure students' opportunities to learn mathematics. Hence, 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-level cognitive processes during instruction.

Mathematical tasks can provide the potential for students to engage in high-level cognitive processes (i.e., problem-solving, reasoning, justification, connecting, or making sense of mathematical ideas) or low-level cognitive processes (i.e., performing rote procedures, memorizing). The level of cognitive demand of a task *as written* reflects the level of cognitive processes required to successfully complete the task. In the "Task Analysis Guide" (TAG), Stein and her colleagues have elaborated four levels of cognitive demand as described in Table 1. The categories of "Doing Mathematics" and "Procedures with Connections" describe tasks with high-level cognitive demands (or "high-level tasks"), and the categories of "Procedures without Connections" and "Memorization" describe tasks with low-level cognitive demands (or "low-level tasks"). The distinction between tasks with high and low cognitive demands makes salient that different types of tasks provide different opportunities for students' learning and place different expectations on students' thinking (Stein et al. 1996). If the tasks students encounter require memorizing facts or practicing procedural computations (i.e., low-level tasks), students are likely to become facile with facts and computational skills. If instructional tasks require students to think, reason, and make sense of mathematical ideas (i.e., high-level tasks, or cognitively challenging tasks), students are likely to become skillful mathematical problem solvers and construct rich understandings of mathematical ideas.

Middle school tasks with different levels of cognitive demand are provided in Table 2. "The Fencing Task" and the task "Using a 10×10 grid…" (Stein et al. 2000) are examples of high-level tasks. "The Fencing Task" requires middle school students to explore multiple possibilities for rectangles with a perimeter of 24 feet, determine which rectangle has the maximum area for the given perimeter, investigate the same situation using a different perimeter, and generalize their findings to apply to any set of rectangles with a fixed perimeter. No solution pathway is suggested by the task, and no well-rehearsed algorithm exists to easily solve the task. Middle school students often begin by using whole-number dimensions, sometimes creating rectangles on grid paper, organizing their information into a table and/or creating a graph. Students are required to conceptualize and determine the perimeter and area of different rectangles, to identify connections between the configuration of the rectangle and the amount of area enclosed, and to justify their conjectures based on the properties and behaviors of the rectangles, table, or graph. These are characteristics of "doing mathematics" tasks, frequently referred to as "open-ended" or "problem-solving" tasks. The "Using a 10×10 grid…" task exemplifies "procedures with connections" tasks by providing a procedure for students to follow, where the suggested procedure leads to a deeper mathematical connection or understanding. In this task, shading in 3/5 of a 10×10 grid allows students to see the connection between 3/5 and .60 ("sixty hundredths") and 60 % (60 out of 100 squares). As an example of a "procedures without connections task," "Martha's Carpeting" (Stein et al. 2000) is straightforward and requires the use of a well-rehearsed algorithm. While "Martha's Carpeting" may be solved by constructing a diagram of the room, nothing in the task requires or supports students to create a representation or develop meaning for area; the task only requires that students produce a single numeric answer. Note that a skillful teacher may set up or implement the task during instruction in ways that provide opportunities for higher-level thinking and reasoning, but the "Martha's Carpeting" task *as written* does not provide such

Table 1 Descriptors of the levels of cognitive demand

Descriptors for high-level tasks from the task analysis guide (Stein et al. 2000)
Doing mathematics tasks
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
Procedures with connections tasks
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, manipulative, 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
Descriptors for low-level tasks from the task analysis guide (Stein et al. 2000)
Procedures without connections tasks
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
Memorization tasks
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

opportunities. Memorization tasks, as illustrated in Table 2, do not require students to use a procedure, but to reproduce memorized knowledge, facts, or formulae.

Though teachers' ability to select high-level instructional tasks is essential in promoting students' learning (NCTM 2000; Simon and Tzur 2004; Stein et al. 2000), teachers typically do not analyze or select instructional tasks by attending to the type or level of thinking that the task can elicit from students (i.e., the level of cognitive demand). Instead,

Table 2 Examples of tasks with different levels of cognitive demand

Doing mathematics	Procedures with connections
High-level cognitive demands	
The Fencing Task (Stein et al. 2000, p. 2)	Stein et al. (2000, p. 13)
Ms. Brown's class will raise rabbits for their spring science fair. They have 24 feet of fencing with which to build a rectangular rabbit pen to keep the rabbits	Using a 10 × 10 grid, determine the decimal and percent equivalents of 3/5
(a) If Ms. Brown's students want their rabbits to have as much room as possible, how long would each of the sides of the pen be?	
(b) How long would each of the sides of the pen be if they had only 16 feet of fencing?	
(c) How would you go about determining the pen with the most room for any amount of fencing? Organize your work so that someone else who reads it will understand it	

Procedures without connections	Memorization
Low-level cognitive demands	
Martha's Carpeting Task (Stein et al. 2000, p. 1)	Stein et al. (2000, p. 13)
Martha was recarpeting her bedroom, which was 15 feet long and 10 feet wide. How many square feet of carpeting will she need to purchase?	What are the decimal and percent equivalents of ½ and ¼?

teachers often rely on instructional materials to provide and sequence instructional tasks, adhering to the tasks in their textbooks or to lists of skills and concepts they need to "cover" (Grouws et al. 2004; Remillard and Bryans 2004). In studies conducted by Stein et al. (1990) and Arbaugh and Brown (2005), teachers categorized tasks with respect to similarities in mathematical content or surface-level features such as "word problems" or "uses a graph." Swafford et al. (1997) found that although teachers were able to create lesson plans that incorporated surface-level features of reform-oriented instruction (such as concluding the lesson with a whole-group discussion), only 3 % of the tasks in the lesson plans had the potential to elicit high-level cognitive demands.

Even when cognitively challenging tasks are selected for instruction, students are not guaranteed opportunities to engage in high-level thinking and reasoning. Maintaining the complexity of high-level tasks is a difficult endeavor (Stigler and Hiebert 2004; Weiss et al. 2003), often shaped by teachers' and students' beliefs about how mathematics is best taught and learned (Lloyd and Wilson 1998; Stein et al. 2007). Teachers and students accustomed to traditional, directive styles of teaching and routinized, procedural tasks experience conflict and discomfort with the struggle that often accompanies high-level tasks. In response to ambiguity or uncertainty on how to proceed, students may disengage with the task or press the teacher for step-by-step instructions, and teachers may reduce high-level demands by breaking the task into less-challenging subtasks or by shifting the focus to correct answers or procedures (Arbaugh et al. 2006; Henningsen and Stein 1997). Furthermore, tasks that have high-level demands as written may not result in high-level thinking and reasoning as implemented in the classroom if they are not appropriately aligned with students' prior knowledge and experiences (i.e., student have too much or too little exposure to similar tasks or to the underlying mathematical ideas). In the TIMSS study, while teachers in the United States used percentages of high-level tasks consistent with the percentages used in many higher-performing countries, the most striking and significant difference between the United States and higher-performing countries in the study was the inability of US teachers to maintain high-level cognitive demands during instruction (Stigler and Hiebert 2004).

Hence, improving students' opportunities to learn mathematics with understanding requires mathematics teachers to select and implement high-level tasks in ways that maintain students' engagement in thinking and reasoning *throughout* an instructional episode. The section that follows describes the design and content of the learning experiences in the "Improving Practice" workshop, intended to catalyze change in teachers' selection and implementation of cognitively challenging instructional tasks.

Enhancing teachers' learning and instructional practices in the ESP workshop

The ESP Project team engaged secondary mathematics teachers in a series of professional learning experiences intended to increase teachers' selection and implementation of instructional tasks that require students to think, reason, and make sense of mathematical ideas. During the year 1 "Improving Practice" workshop (2004–2005), teachers attended six one-day sessions held on Saturdays. Table 3 provides an overview of the specific activities in the "Improving Practice" workshop. The book, "*Implementing Standards-Based Mathematics Instruction: A Casebook for Professional Development*" (Stein et al. 2000) served as the "textbook" (hereafter referred to as the *Casebook*) and provided the conceptual framework.

The "Improving Practice" workshop utilized a practice-based approach to teachers' learning, where professional learning activities were grounded in the actual work of teaching and facilitators modeled the pedagogy of good instruction (Ball and Cohen 1999; Smith 2001). A central sequence of activities in the workshop was to engage teachers in solving a cognitively challenging mathematical task, analyze the cognitive demands of the task, and reflect upon instructional artifacts (i.e., students' work) or an instructional episode (i.e., narrative or video cases) of a teacher using the task in a mathematics lesson. Hence, the ESP workshop provided frequent opportunities for teachers to increase their knowledge of mathematics, knowledge of effective mathematics pedagogy, and knowledge of students as learners of mathematics as described in Shulman's (1986) seminal paper. For example, solving challenging tasks allowed teachers to deepen their mathematical knowledge and gain first-hand experience as "learners" in reform-oriented mathematics lessons. By reflecting on how the experience contributed to their own learning, teachers come 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 students' learning (Borasi et al. 1999; Farmer et al. 2003). Solving challenging tasks, thus, provides opportunities to enhance teachers' *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 of different representations of mathematical concepts) (Hill et al. 2008).

After solving a task and reflecting on their experiences as learners, teachers also analyzed the cognitive demands of the task. The discussion of cognitive demands initiated in Session 1 consisted of a comparison between "Martha's Carpeting Task" and the "Fencing Task" (Stein et al. 2000; see Table 2), two tasks similar in mathematical content (i.e., the mathematical idea of area is present in each task, though in different ways), but very different in cognitive demand (as described previously). Following this comparison,

Table 3 Activities in the ESP year 1 "Improving Practice" workshop

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 use cases?	Case stories II: storytelling through student work	Case stories III: How did assessing & advancing questions influence the enactment of the task?
Solving Martha's carpeting task & the fencing task	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		Planning the "Sharing and Discussing" Phase of a Lesson
Comparing Martha's carpeting task & the fencing task	Reading & discussing the case of Ron Castleman	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	Introducing the "Thinking Through a Lesson Protocol" (TTLP)
Categorizing mathematical tasks	The factors and patterns of maintenance & decline	Reading & discussing the case of Monique butler	Analyzing student work on the extend pattern of tiles task	Planning a whole-group discussion	Solving the double the carpet task
Data collection, paperwork	Data collection, paperwork	Connecting to own teaching: discuss factors that influenced your lesson			Data collection, paperwork
Assignments					
Identify a task from your data collection that you would like to adapt or improve in some way	Plan, teach, and reflect on a lesson involving a high-level task; identify 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?	Plan, teach, and reflect on a lesson involving a high-level task. Bring in student work	Plan, teach and reflect on a lesson involving a high-level task. Bring in a list of questions and student work	Plan, teach and reflect on a lesson involving a high-level task. Use the TTLP to plan your lesson

teachers analyzed the level of cognitive demand of a set of middle school mathematics tasks (Smith et al. 2004), discussed their individual criteria for categorizing tasks as "High-Level" or "Low-Level," and collectively constructed a set of criteria for high- versus low-level cognitive demands. In Session 2, teachers were introduced to the Task Analysis Guide (TAG) (see Table 1), and their analysis of the cognitive demands of mathematical tasks progressed from a dichotomous categorization of high level versus 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 in every session except Session 5. When categorizing the cognitive demands of a *task as written*, teachers consider the level of cognitive processes required to solve the task, with the assumption that the task is instructionally appropriate for the context and students with which it is being used. (If the task was not appropriate, this would be evident in, and analyzed as a part of, the *task implementation* as students worked on the task during instruction.)

Other tools and ideas from the *Casebook* (e.g., the Mathematical Tasks Framework) supported teachers to analyze task implementation in artifacts and episodes of instruction, specifically, in sets of students' work, narrative cases (in the *Casebook* and in Smith et al. 2005), and videotaped cases of instruction featuring the cognitively challenging tasks teachers had engaged in solving. Discussions of instructional cases and artifacts provided opportunities for teachers to examine mathematics pedagogy that challenged traditional views of effective teaching and learning, without initially having to focus on their own practice (Kazemi and Franke 2004; Sherin and van Es 2005). At the end of each session, teachers were given assignments, or scaffolded field experiences (SFE's) (Borasi and Fonzi 2002), to apply the ideas from the professional development workshop into their own classrooms (see the bottom row of Table 3). A significant portion of Sessions 3–5 involved teachers sharing their personal experiences using cognitively challenging tasks in their own classrooms, or telling their own "case stories" (Hughes et al. 2008). These experiences were collectively intended to enhance teachers' pedagogical content knowledge, specifically, teachers' "knowledge of content and teaching" (i.e., knowledge of the cognitive demands of mathematical tasks and teachers' decisions in selecting instructional tasks) and "knowledge of students and content" (i.e., considering what questions to ask to assess and advance student's thinking, or considering how to select and order presentations of students' work) (Hill et al. 2008; Shulman 1986).

In the conceptual model presented in Fig. 1, enhancing teachers' knowledge is situated as the mediator between professional learning experiences and changes in teachers' instructional practices. Providing teachers with opportunities to solve challenging mathematical tasks and to evaluate the task demands (as written and as implemented during instruction) was anticipated to enhance teachers' knowledge of cognitive demands and of the influence of cognitively challenging tasks on students' learning. As described previously, an analysis of instructional tasks, collections of students' written work, and lesson observations from ESP teachers' classrooms indicated a significant increase in teachers' use of high-level instructional tasks and in teachers' ability to implement high-level tasks (Boston 2006; Boston and Smith 2009). Given these important changes in teachers' instructional practices, this paper analyzes what teachers learned about the cognitive demands of mathematical tasks throughout their participation in the "Improving Practice" workshop (research question 1) that may have catalyzed change in their selection and implementation of cognitively challenging instructional tasks. This investigation also explores the "arrows" connecting the professional learning experiences and teachers' knowledge, that is, where in the workshop did teachers have opportunities to enhance their knowledge of cognitive demands (research question 2). The methods for exploring these questions are presented in the next section.

Methodology

This investigation explores changes in teachers' knowledge of the cognitive demands of mathematical tasks following their participation in the ESP "Improving Practice" workshop throughout the 2004–2005 school year and how those changes connect back to teachers' experiences in the workshop. This section describes the subjects, data sources, and analyses in the study.

Subjects

Nineteen secondary mathematics teachers participated in this study, ranging from three to 30 years of teaching experience with an average of 8.5 years in the classroom. At the time of their participation in the study, ten teachers were teaching middle school mathematics (grades 6–8) and nine were teaching high school mathematics (grades 9–12). The teachers represented ten school districts, all within the city limits or bordering the same mid-sized city. School demographics included a large urban school district, a mid-sized affluent school district, and several small school districts in middle- to working-class neighborhoods. 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. A group consisting of ten teachers (five each from two schools in the region) was selected to compare whether the knowledge and instructional practices of ESP teachers following their participation in the workshop would differ from the knowledge and instructional practices of teachers who did not participate in the ESP project (and were not provided with professional development experiences). The contrast group teachers and their schools reflected the variety in demographics, type of curriculum, and participation in professional development of the ESP group (Boston 2006).

Data on teachers' knowledge: The task sort

ESP teachers and contrast group teachers completed a written-response card sort, with each card (i.e., half-sheet of paper) containing one mathematical task, as an assessment of their knowledge of the cognitive demands of mathematical tasks. The design, use, and analysis of the "task sort" was informed by prior research in which card sorts were used to elicit teachers' content knowledge (Stein et al. 1990) and to assess teachers' ability to categorize mathematical tasks (Arbaugh and Brown 2005). ESP teachers completed the task sort as an individual written activity during the first session in October 2004 and during the final session in May 2005. Teachers in the contrast group completed the task sort in May 2005, after having no participation in ESP or related professional development opportunities throughout the 2004–2005 school year. Participants were given 16 cards containing a mathematical task and prompted to: (1) rate the task as High, Low, or Not Sure and (2) provide a rationale for the rating. After rating all of the tasks, participants were asked to list their criteria for high- and low-level tasks. A sample task-sort card is provided in Fig. 2.

The middle school task set created by Smith et al. (2004) was used as the source of tasks for the task-sort analysis. This task sort was designed to allow subjects to categorize tasks

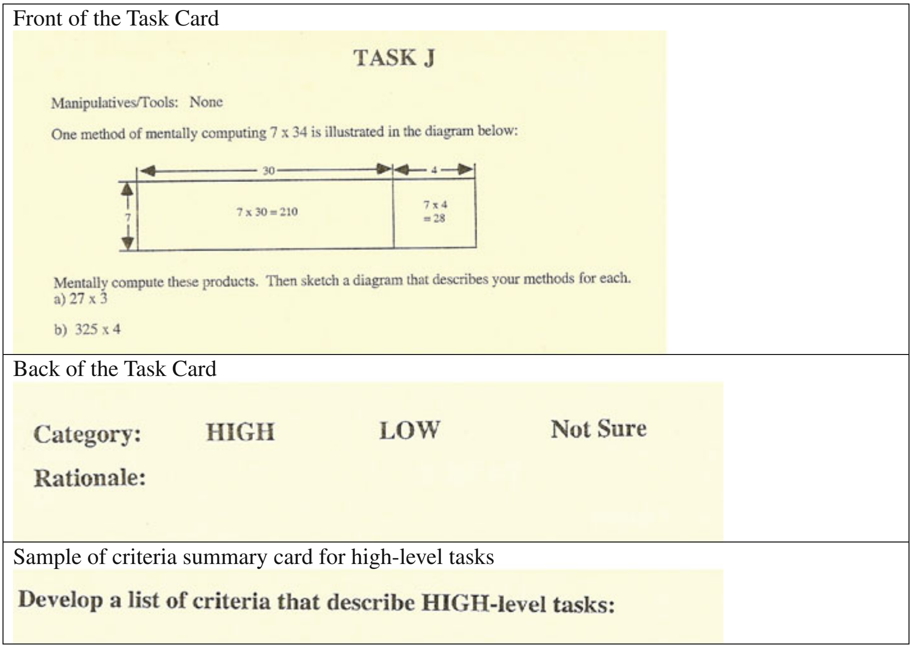

Fig. 2 Sample task cards from the task sort

based on several possible criteria, such as mathematical content, use of representations (i.e., diagrams or graphs) or manipulatives, use of a context, or requirement of an explanation. 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 set contains 16 tasks total: two "Memorization" tasks, four "Procedures without Connections" tasks, five "Procedures with Connections" tasks, and five "Doing Mathematics" tasks.

Grover's (1989) framework for scoring teachers' verbal responses to open-ended interview questions informed the coding of teachers' task-sort responses in this investigation. Task-sort responses were scored for a total of 38 points using the following dimensions: (1) one point per task for identifying the correct level of cognitive demand; (2) one point per task for providing a rationale consistent with the cognitive demands of the task (i.e., consistent with descriptors in the Task Analysis Guide [TAG], Table 1); and (3) overall criteria consistent with TAG descriptors for high-level tasks (0–3 points) and for low-level tasks (0–3 points) (Boston 2006). All 48 task-sort responses (19 pre-test, 19 post-test, and 10 contrast subjects) were scored by the primary investigator (author), and a double-blind, stratified random sample of 10 of the responses (20.8 %; four pre-test, four post-test, two contrast group) was scored by a trained rater to check the consistency of the scores and the reliability of the scoring system. Agreement of 92.9 % was reached on the item-by-item scores of the two raters.

To examine whether any factors external to the professional development may have contributed to changes in teachers' knowledge of the cognitive demands of mathematical tasks, task-sort scores were compared between teachers using standards-based and traditional curricula and between ESP and contrast group teachers. These comparisons determined: (1) whether exposure to and use of a standards-based curriculum (i.e., curricula identified as "exemplary or promising" by the US Department of Education (USDE 1999), rated highly by the American Association for the Advancement of Science (AAAS 2000), and/or developed under the auspices of a National Science Foundation grant) in the classroom would provide teachers with more knowledge or growth in knowledge of cognitive demands than the use of a traditional curricula and (2) whether the activity of teaching mathematics over a school year, without the intervention of a professional development workshop, would increase teachers' ability to identify and describe tasks with high and low levels of cognitive demand.

Descriptive statistics were used to assess the pre- and post-workshop task-sort scores from ESP teachers and the task-sort scores from contrast group teachers. Non-parametric, one-tailed statistical tests with a significance level of $p < .01$ were used to identify increases in teachers' pre- versus post-workshop task-sort scores and differences between ESP teachers and the contrast group. A two-way ANOVA, with "curriculum type" and "time" as the grouping variables, was conducted to determine whether the use of a reform versus traditional curriculum in teachers' classrooms influenced their knowledge of the cognitive demands of mathematical tasks before and after their participation in ESP.

This study utilizes a mixed methods approach, aligned with current calls for more rigorous professional development research (Desimone 2009; Scher and O'Reilly 2009). When statistically significant differences were found, the nature of the differences was analyzed and described qualitatively, thus affording precise descriptions of *what* teachers learned and *where* they had opportunities to learn it throughout the professional development sessions. Task-sort responses were coded qualitatively for: (1) use of specific categories from the TAG (i.e., memorization, procedures without connections, procedures with connections, doing mathematics); (2) statements inconsistent with descriptors in the TAG (i.e., "low-level tasks contain diagrams"); and (3) prominent words or phrases used by at least nine of 19 ESP teachers on the post-workshop task-sort responses. Emergent or prominent language on the post-workshop task-sort response was considered to represent ideas teachers developed or refined during the workshop, particularly if this language was not present on the pre-workshop responses nor in the responses from the contrast group. The qualitative codes are provided with the results in Table 7. Reliability for qualitative data was determined by providing the coding categories to a graduate research assistant not associated with the ESP project for double-blind coding of all 48 task-sort responses. In 126 of 133 instances (97 %), the graduate student identified the same teachers in each category as identified by the Principal Investigator. The Principal Investigator resolved discrepancies by reviewing the specific task-sort response for evidence of the teachers' use of language reflective of that category.

Data on teachers' participation in the ESP professional development workshop

Videotapes of the six professional development sessions in the year 1 "Improving Practice" workshop (2004–2005) were reviewed by the Principal Investigator (author), and segments of the videotapes were identified as providing opportunities for teachers' learning. Codes included: (1) opportunities to engage in learning about the cognitive demands of mathematical tasks; (2) specific use of the Task Analysis Guide (Table 1); (3) conversations reflecting the prominent language used on ESP teachers' post-workshop task-sort responses (Table 4); and (4) conversations containing ideas about the cognitive

Table 4 Descriptive statistics on task-sort scores

	n	Pre-workshop Mean (SD)	Post-workshop Mean (SD)
All ESP teachers	19	24.21 (5.75)	28.74 (4.12)*
ESP teachers using reform curricula	10	24.10 (5.78)	28.00 (5.08)*
ESP teachers using traditional curricula	9	24.33 (6.06)	29.56 (2.79)*
ESP teachers with no prior exposure to task sort	14	22.86 (5.99)	29.00 (5.25)*
Contrast group teachers	10	NA	17.60 (6.13)**

* Significantly higher than re-workshop score at $p < .01$

** Significantly lower than post-workshop scores for all other categories at $p < .01$

demands of mathematical tasks that were not prominent in teachers' language on the post-workshop task-sort responses. A graduate research assistant not associated with the ESP project reviewed all of the videotapes using these codes and identified 48 of the 51 (94 %) instances identified by the Principal Investigator and six instances not previously identified. Differences were reviewed by the Principal Investigator, with five of the new instances confirmed for inclusion in the results. The graduate assistant also performed a "reverse coding" of the video, that is, viewed the video for prominent ideas that arose during teachers' discussions of solving and analyzing mathematical tasks to identify any ideas prominent in the workshop that did not arise on ESP teachers' post-workshop task-sort responses. Finally, self-report data (i.e., teachers' statements transcribed from videotape; teachers' writings during the professional development sessions) were utilized to provide instantiations of the nature of changes in teachers' knowledge of cognitive demands throughout their participation in the professional development sessions.

In summary, changes in teachers' knowledge (research question 1) were assessed using a task-sort instrument, and these changes were connected to teachers' experiences in the ESP "Improving Practice" workshop (research question 2) through video analysis and teachers' self-reports.

Results

Analyses of the task-sort responses, videotapes, and artifacts from the professional development workshop identify what new ideas ESP teachers appeared to have learned, how their newly acquired knowledge differed from teachers in the contrast group, and when (throughout the "Improving Practice" workshop) ESP teachers had opportunities to acquire these new ideas.

Changes in teachers' knowledge

Table 4 provides the mean task-sort scores used in the analyses. Scores on the pre-workshop task sort ranged from 13 to 32 points (out of a possible 38), 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 was significant ($z = 3.15$; $p < .001$ [one-tailed]). A two-way ANOVA (with time and curriculum as the grouping variables) indicated that the difference in means between ESP teachers using reform-oriented versus traditional curricula was not significant ($F = .192$; $p = .667$) and that teachers using one type of curricula did not increase in scores significantly more than teachers using the other type of curricula ($F = .324$; $p = .577$). ANOVA results also confirm the increase in task-sort scores over time ($F = 15.424$; $p < .001$).

ESP teachers' pre-and post-workshop task-sort scores were compared to the task-sort scores of the contrast group using Mann–Whitney tests. The task-sort scores from the 10 contrast group teachers ranged from eight to 26 points, with a mean of 17.6. Personal communication during the first session indicated that five teachers had engaged in a "task sorting" activity prior to their participation in ESP, and the task-sort scores from these teachers were excluded from comparisons with the contrast group. Prior to their participation in the workshop, the task-sort scores of the 14 ESP teachers with no prior exposure to the task sort were similar to the task-sort scores of the contrast group ($z = 1.55$; $p = .12$ [two-tailed]). Following their participation in the workshop, teachers' post-workshop task-sort scores were significantly higher than those of the contrast group for the 14 ESP teachers with no prior exposure ($z = 3.63$; $p < .001$ [one-tailed]) and for the 19 ESP teachers overall ($z = 3.95$; $p < .001$ [one-tailed]).

Teachers' specific learning about cognitive demands

ESP teachers' pre- and post-workshop task-sort responses were analyzed qualitatively to determine the nature of teachers' learning, specifically, whether teachers improved their ability to identify the level of cognitive demand and/or to provide criteria and describe the features of (and opportunities for students' learning provided by) high- and low-level tasks.

Identifying the level of cognitive demand

Data in Tables 5 and 6 illustrate the nature of changes in teachers' task-sort responses over time, by number of teachers and by number of task classifications, respectively. Significant increases in teachers' task-sort scores cannot be attributed to an increased ability to identify high-level tasks. Teachers were successful at classifying "Doing Mathematics" tasks as high level on both the pre- and post-workshop task sort. When teachers incorrectly classified a doing mathematics task as having low-level demands on the post-workshop task sort, their rationales indicated that: (1) the task did not require an explanation; (2) the task did not connect to a real-world context; or (3) the task did not require mathematical thinking. On the contrary, teachers had difficulty categorizing "Procedures with Connections" tasks as high-level tasks on both the pre- and post-workshop task sort. At both pre- and post-workshop, procedures with connections tasks were categorized incorrectly three times as often as doing mathematics tasks. On the post-workshop task sort, the predominant rationale teachers provided for classifying a procedures with connections task as low level was the presence of a stated or implied procedure or "steps" for solving the task (i.e., procedures with connections were classified as "Procedures without Connections" tasks), overlooking the opportunities for developing mathematical connections and understanding embedded in the task. For example, "Task J" in Fig. 2 provides a procedure to solve the task; however, in following the procedure (using a rectangular area model to illustrate their mental multiplication strategy), students gain insight into multi-digit multiplication and the distributive property. The procedure provides opportunities to make mathematical connections, and the procedure itself is not the mathematical goal of the task.

ESP teachers were proficient in identifying low-level tasks on the pre-workshop task sort, and this ability improved slightly over time. When procedures without connections

Table 5 Analysis of the task-sort responses by number (%) of teachers

Level of cognitive demand	# of Tasks	Number of teachers incorrectly classifying a task			Number of teachers providing criteria synonymous with TAG		
		ESP pre-workshop ($n = 19$)	ESP post-workshop ($n = 19$)	Contrast ($n = 10$)	ESP pre-workshop ($n = 19$)	ESP post-workshop ($n = 19$)	Contrast ($n = 10$)
High level							
Doing mathematics	5	10 (53 %)	9 (47 %)	9 (90 %)	14 (74 %)	19 (100 %)	4 (40 %)
Procedures with connections	5	17 (89 %)	15 (79 %)	10 (100 %)	5 (26 %)	10 (53 %)	0 (0 %)
Low-level							
Procedures without connections	4	12 (63 %)	7 (37 %)	6 (60 %)	14 (74 %)	19 (100 %)	6 (60 %)
Memorization	2	4 (21 %)	0 (0 %)	4 (40 %)	10 (53 %)	19 (100 %)	3 (30 %)

Table 6 Analysis of the task-sort responses by number (%) of incorrect classifications

Level of cognitive demand	# of Tasks	Total number of classifications[a]		Number of incorrect classifications		
		ESP (19 teachers)	Contrast (10 teachers)	ESP pre-workshop (19 teachers)	ESP post-workshop (19 teachers)	Contrast (10 teachers)
High level						
Doing mathematics	5	95	50	16 (17 %)	15 (16 %)	16 (32 %)
Procedures with connections	5	95	50	49 (52 %)	47 (49.5 %)	31 (62 %)
Low-level						
Procedures without connections	4	76	40	15 (20 %)	7 (9 %)	8 (20 %)
Memorization	2	38	20	4 (11 %)	0 (0 %)	5 (25 %)

[a] Total number of classifications are determined by multiplying the number of tasks at that level by the number of teachers

tasks were classified as high level, teachers' rationales indicated that the task required an explanation or contained a real-world context. No teachers misclassified a "Memorization" task at post-test.

Providing rationales for task levels

As indicated in Table 5, teachers' ability to provide rationales for high- and low-level cognitive demands improved over time. In their criteria for tasks on the post-test, three results are noteworthy. First, more teachers used the specific task labels from the TAG (i.e., doing mathematics, procedures with connections, procedures without connections, and memorization) for all four task categories. Second, all 19 teachers used language synonymous with descriptors from the TAG (i.e., describing doing mathematics tasks as "open-ended" or "problem-solving" or procedures without connections tasks as "procedural," "computation," or "basic skills") for every task category except procedures with connections. Third, while six teachers listed criteria contradictory to features of high- or low-level tasks on the pre-test (i.e., a task is low level if it "contains a diagram" or "uses manipulatives"; a task is high level if it "is beyond students' reach" or "difficult for my students"), no teachers provided contradictory criteria on the post-test.

Use of new language

ESP teachers used specific terminology to describe the features of high- and low-level tasks on their post-workshop task-sort responses. In many cases, this terminology was not prominent on their pre-workshop responses, nor was it evident on contrast teachers' task-sort responses. As displayed in Table 7, language that emerged as prominent on the post-workshop task-sort responses included: *representations*, *generalizations/generalize*, *connections*, and *multiple strategies/open-ended*. Other terminology prominent on the pre-test was used by an even greater number of ESP teachers on the post-test, for example, *explanations*, *procedures/procedural*, and phrases synonymous with "procedures without connections" (i.e., *computational*, *drill, skills practice*).

Comparisons to the contrast group

Qualitative comparisons also illuminated interesting similarities and differences between the task-sort responses of ESP teachers and contrast group, provided in Table 7. Similar to ESP teachers, contrast group teachers experienced the most difficulty identifying and describing the characteristics of procedures with connections tasks. Eight contrast group

Table 7 Comparison of task-sort responses between ESP teachers and contrast group teachers

	ESP pre-workshop ($n = 19$)	ESP post-workshop ($n = 19$)	Contrast group ($n = 10$)
Number of teachers using specific level of cognitive demand			
Doing mathematics	0 (0 %)	1 (5 %)	0 (0 %)
Procedures with connections	2 (11 %)	9 (47 %)	0 (0 %)
Procedures without connections	9 (47 %)	15 (79 %)	0 (0 %)
Memorization	10 (53 %)	19 (100 %)	3 (30 %)
Number of teachers using specific terminology			
Representations	3 (16 %)	9 (47 %)	3 (30 %)
Explanations; explain	13 (68 %)	17 (89 %)	1 (10 %)
Multiple strategies; open-ended	7 (37 %)	15 (79 %)	1 (10 %)
Generalization; generalize	4 (21 %)	9 (47 %)	0 (0 %)
Connections	2 (11 %)	12 (63 %)	0 (0 %)
Procedural; procedures; or synonymous (calculations; computations; drill; basic skills practice)	9 (47 %) 14 (74 %)	15 (79 %) 17 (89 %)	0 (0 %) 4 (40 %)
Number of teachers making statements contradictory to task analysis guide	6 (32 %)	0 (0 %)	8 (80 %)

teachers (80 %) provided criteria contradictory to characteristics of high- or low-level tasks, compared to six pre-workshop ESP teachers (32 %) and 0 post-workshop ESP teachers. In the 10 contradictory statements, seven contrast group teachers identified high-level tasks as "difficult" and focused on whether students would "struggle" or "have problems" with the task; three indicated that low-level tasks contain "visual aids" or "diagrams." Contrast group teachers were less likely to identify the specific levels of cognitive demand of mathematical tasks or provide criteria that included ideas synonymous with the TAG. Rarely did contrast group teachers use the emergent or prominent language identified on the ESP teachers' post-workshop task-sort responses, with the exception of synonyms of procedures without connections (i.e., *calculations*), identified by four of 10 (40 %) contrast group teachers, and *representations*, identified by three of 10 (30 %) contrast group teachers. *Explanations* and *multiple strategies, open-ended,* were each identified by one contrast group teacher, and none mentioned *generalizations, connections,* or *procedures*.

Connecting teachers' learning to the ESP Workshop

Throughout the "Improving Practice" workshop, teachers often engaged in discussions about the cognitive demands of mathematical tasks. Analysis of the videotaped professional development sessions identified: (1) when ESP teachers had opportunities to increase their knowledge of the cognitive demands of mathematical tasks and (2) when ESP teachers had opportunities to develop the ideas associated with the terminology that emerged or was prominent on the post-workshop task-sort responses. Table 8 provides an overview of this analysis.

Opportunities to learn about cognitive demands

Discussions about the cognitive demands of mathematical tasks occurred in five of the six sessions, as teachers discussed the cognitive demands of the tasks they had engaged in solving (see Table 3). Opportunities for teachers to compare and categorize tasks occurred during the first two sessions, as teachers compared "Martha's Carpeting" and the "Fencing Task" (Table 2), engaged in a task-sorting activity, were introduced to the TAG and levels of cognitive demand (Table 1), and used the TAG to categorize tasks. Teachers' experiences solving tasks as learners (in each session except Session 5) also appear to have provided opportunities to increase their knowledge of cognitive demands, as many of the ideas represented in teachers' emergent and prominent language on the post-workshop task sort arose during discussions of and reflections on their own mathematical work.

Emergent and prominent language

How do teachers' learning opportunities throughout the "Improving Practice" workshop compared to the changes identified on the pre- and post-workshop task-sort responses? Table 8 identifies where the specific criteria for high- and low-level tasks prominent on teachers' post-test responses arose during discussions throughout the workshop. Interestingly, "use of a diagram" also emerged when coding the discussions, but this criterion was not present on the post-workshop task-sort responses. Also notable was that teachers consistently associated the presence of a procedure as a feature of low-level tasks and the request for an explanation as a necessary criterion of high-level tasks.

Table 8 Ideas about the cognitive demands of mathematical tasks that arose during discussions in the ESP workshop

Session	Opportunities to discuss level of cognitive demands	Specific use of TAG	Multiple solution strategies	Multiple representations	Generalizations	Connections	Presence of Procedure	Explanation	Use of diagrams
1	Solving "Fencing" task		X	X	X	a, b	X	X	
	Comparing Tasks		X	X	X	a, b, c			
	Categorizing Tasks		X	X	X	c			
2	Introduction of TAG	X	X			c	X	X	X
	Solving "Linking" task	X	X			c			X
	Level of "Linking" task								
3	Solving "Algebra Tiles" task	X	X	X		b, c			X
	Level of Algebra Tiles task								
4	Solving "Pattern of Tiles" task	X	X	X	X	a, b, c	X	X	X
	Level of "Pattern of Tiles" task		X	X		c			X
6	Solving "Double the Carpet" task	X	X	X	X	a, b, c		X	X
	Level of "Double the Carpet" task								

a Indicates connections between strategies

b Indicates connections between representations

c Indicates connections between mathematical concepts

Emergent and prominent criteria were often explicitly modeled by the facilitators during discussions of teachers' work on mathematical tasks (i.e., "Were you surprised by all of the different strategies?" [video transcript, Session 2]; "What is different about Iris and Randy's strategy?" [video transcript, Session 4]; "How does the equation connect to the diagram?" [video transcript, Session 6]). In the five sessions where teachers were asked to identify the cognitive demands of the tasks they had engaged in solving, teachers explicitly identified high-level features of their own work on the task as characteristics that gave the task high-level cognitive demands. During Session 1, teachers presented and discussed multiple solution strategies and representations for solving the "Fencing Task" (Table 2). The facilitator made explicit moves to foster connections between strategies and between representations (i.e., "Do you see any connections between Randy and Dave's solutions?"; "What is it about the table that gives you a clue about the graph?" [video transcript, Session 1]). During the comparison of the "Fencing Task" and "Martha's Carpeting Task" (Table 2), teachers identified features that were prominent in their own work on the task (i.e., multiple strategies, multiple representations, and connections between strategies and between representations) as characteristics that made the "Fencing Task" *different* from "Martha's Carpeting." Comments from two participants during the comparison of the tasks illustrate that teachers were drawing on their experiences in solving the tasks as learners (video transcript, Session 1):

> 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.

Similarly, while solving a high-level task during Session 2, teachers were provided with resources to enable them to create a variety of strategies and were prompted to use a 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' mathematical work in Session 4 focused on using diagrams and forming generalizations, and these criteria also emerged during the discussion of the level of cognitive demand of the task.

Discussion

At the end of the "Improving Practice" workshop, ESP teachers significantly increased their knowledge of the cognitive demands of mathematical tasks and had significantly higher knowledge than teachers in the contrast group. This discussion highlights the changes in teachers' knowledge, poses a hypothesis for connecting teachers' learning to the observed changes in teachers' instructional practices, and suggests improvements to the "Improving Practice" workshop.

Increase in teachers' knowledge of cognitive demands

The nature of improvements in ESP teachers' pre- to post-workshop task-sort responses help substantiate that the increases in 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) and provide evidence that teachers did not simply learn the "correct" answers throughout their participation in ESP. The improvements occurred in teachers' *criteria* and *rationales* for describing high- and low-level tasks. This finding is consistent with prior research where notable differences existed between the *type of criteria* provided by novice and expert teachers (Stein et al. 1990) and in teachers' pre- and post-workshop responses (Arbaugh and Brown 2005). The responses from novices and pre-workshop teachers consisted of surface-level features, while response from experts and post-workshop teachers identified aspects of the tasks that provided opportunities for understanding, sense-making, and students' thinking and actions required to solve the tasks. Similarly, ESP teachers' post-workshop task-sort responses reflected an enhanced knowledge of the characteristics of mathematical tasks that influence students' opportunities for high-level thinking and reasoning (i.e., generalizations, representations, connections). The emergent and prominent language used in teachers' post-workshop task-sort responses provides evidence that teachers became more aware of how high-level tasks support students' learning. Engaging with high-level tasks as learners also appears to have allowed teachers to implicitly attend to features and characteristics of tasks that provide opportunities for high-level thinking and reasoning.

Specific overgeneralizations can account for the misclassification of tasks on the post-test. First, teachers often overlooked the underlying mathematical concepts or connections embedded in high-level tasks. In procedures with connections tasks, half of the ESP teachers persisted in identifying the presence of a procedure (rather than the opportunities for connections) as the feature that determined the level of cognitive demand ("procedures = low-level"). In the most frequently missed doing mathematics task (a qualitative graphing task that requires students to write a story relating speed and time), the six teachers who consider the task to be low level indicated that the task did not require mathematical thinking. Second, in other doing mathematics tasks, a small set of teachers focused on features of the task that appeared to be "missing," such as a real-world context or a prompt for an explanation, and appeared to consider both as necessary conditions for high-level demands. The overgeneralizations that "explanation = high level" and "context = high-level" also account for misclassifications of procedures without connections tasks (at post). For example, a contextual task that required students to apply a well-rehearsed procedure and to "explain the process you used to find the sale price" was regarded as high level. The importance of pressing students to explain their thinking and reasoning was a prominent theme across the ESP workshop, in teachers' rationale for high-level tasks and in discussions about effective questioning and supporting students' work on high-level tasks (see Fig. 2); hence, the workshop may have inadvertently fostered this overgeneralization. In contrast to the pre-test, none of the overgeneralizations that appear on the post-test were contradictory to ideas or descriptors in the TAG. Teachers no longer rated the level of demand based on the perceived difficulty of the mathematical content or skills (i.e., considering long division with decimals high level because "this is a difficult skill for my students," rating a problem-solving task as low level because the underlying mathematics was "easy to solve"). The idea that a diagram or manipulative characterized low-level tasks appears to have dissipated as well, perhaps due to teachers' exposure to high-level tasks that incorporated diagrams or manipulatives throughout the workshop. Conversely, however, the presence of a diagram was not a prominent theme on teachers' post-responses as a criterion for high-level tasks.

Differences between ESP teachers' pre- and post-workshop task-sort responses and between ESP and contrast teachers' responses indicate that ESP teachers learned to characterize tasks with high- and low-level cognitive demands using ideas in the Task Analysis Guide (Table 1) and other ideas made salient in the workshop. At the close of the workshop, ESP teachers used language for describing the cognitive demands of mathematical tasks different from language they had used prior to the workshop and different from language used by teachers who had not participated in the workshop.

Connecting teachers' learning to changes in their instructional practices

ESP teachers improved their ability to identify aspects of tasks that provide opportunities for different levels and types of student thinking (i.e., cognitive demands). One plausible hypothesis is that teachers selected significantly more high-level tasks for instruction after their experiences in the workshop because they learned to attend to and value the opportunities for students' learning embodied in such tasks. This hypothesis is consistent with Remillard and Bryans' notion of 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). Teachers' conceptions "act as critical filters" (Lloyd and Wilson 1998, p. 250) that create differences, and often vast disparities, between teachers' interpretation and implementation of the curriculum and the intentions of the curriculum developers (Remillard and Bryans 2004; Stein et al. 2007). ESP teachers increased their knowledge of the cognitive demands of mathematical tasks and their awareness of how high-level tasks support students' learning, which together changed teachers' orientation toward their curricula (reform or traditional) in ways that supported the selection of high-level instructional tasks in their own classrooms.

Improvements to the ESP "Improving Practice" workshop

Though effective in transforming teachers' knowledge and instructional practices, the ESP "Improving Practice" workshop could be improved in ways that would further enhance teachers' knowledge of the cognitive demands of instructional tasks. The category of *procedures with connections* was the most frequently missed on the task sort, with only half of the teachers identifying this category in their criteria for high-level tasks. A discussion during Session 6 elucidates this finding. Participants were commenting on the "Thinking Through a Lesson Protocol" (Smith et al. 2008), 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 (video transcript, Session 6):

> 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 [*pattern-generalizing*] task and you have days where you learn FOIL? 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.

The discussion continued, lasting almost 14 min, 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.

Interestingly, ten teachers referred to this discussion in the session evaluation, and three teachers referred to it in their post-workshop interview approximately 1 month later. This indicates that, in the last session of the workshop, teachers were still wrestling with the idea of "procedures with connections." The significance of this discussion to a majority of the ESP teachers suggests that connecting mathematical procedures to meaning and understanding was not adequately addressed within the workshop. Further research would be needed to determine whether this deficiency might be improved through professional learning activities that enhance teachers' content knowledge (i.e., understanding the underlying mathematics and possible connections in procedures with connections tasks) or pedagogical content knowledge (i.e., considering how to revise procedural tasks to provide opportunities for connections and sense-making).

Conclusion

The results of this investigation are important for several reasons. Most significantly, teachers in the study improved their knowledge and instructional practices along dimensions of teaching that have been linked to increases in students' opportunities for learning. Teachers increased their knowledge of the ways in which cognitively challenging tasks support students' learning, and subsequently selected more cognitively challenging tasks and improved their ability to maintain high-level demands during implementation.

Many of the findings have broad implications beyond the specifics of this project. First, teachers' experiences in solving mathematical tasks appear to have greatly influenced their learning. New ideas prominent on ESP teachers' post-workshop task-sort responses, and different from the responses of the contrast group, were remarkably consistent with ideas that arose during discussions of teachers' work in solving challenging tasks as learners. Second, teachers' difficulty seeing the potential of *procedures with connections tasks* to support students' learning, and instead focusing only the procedure, might help explain the difficulty of implementing tasks in ways that provide students with opportunities to make mathematical connections identified in large-scale studies. Third, the use of a mixed methods approach facilitated connections between teachers' experiences in the ESP workshop and gains in their knowledge of cognitive demands. In this investigation, the professional development activities, goals for teachers' learning, and the methodology used to assess teachers' learning all focused on teachers' selection and implementation of high-level tasks. Though not establishing causal links, the strong connections between changes in teachers' knowledge and their experiences in the "Improving Practice" workshop provide indications that learning occurred during the workshop, and this learning may have influenced subsequent changes in teachers' classroom practices identified by Boston and Smith (2009). Ideas for future work include replicating the ESP workshop at scale, with larger groups of teachers, to analyze the extent to which this type of "task-based" professional development can invoke systemic change.

Acknowledgments The work described in this article was conducted under the auspices of the ESP Project, through a grant from the National Science Foundation (DUE-0301962), and is based on dissertation research conducted at the University of Pittsburgh by Melissa D. Boston under the directorship of Margaret S. Smith. Any opinions expressed herein are those of the authors and do not necessarily represent the views of the Foundation. The author would like to thank Brian Molinero, Duquesne University, for his assistance with reliability coding for the task-sort responses and videotaped ESP sessions. The author would also like to thank Margaret Smith, Fran Arbaugh, and three anonymous reviewers for their thorough and thoughtful feedback on earlier versions of this manuscript.

References

American Association for the Advancement of Science (AAAS). (2000). *Middle grades mathematics textbooks: A benchmarks-based evaluation.* Retrieved October 7, 2011, from http://wwwproject2061.org/publications/textbook/mgmth/report/default.htm.

Arbaugh, F., & Brown, C. A. (2005). Analyzing mathematical tasks: A catalyst for change? *Journal of Mathematics Teacher Education, 8*, 499–536.

Arbaugh, F., Lannin, J., Jones, D. L., & Park-Rogers, M. (2006). Examining instructional practices in Core-plus lessons: Implications for professional development. *Journal of Mathematics Teacher Education, 9*, 517–550.

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.

Borasi, R., & Fonzi, J. (2002). Engaging in scaffolded instructional innovation. *Foundations: Professional Development that Supports School Reform* (pp. 83–98). Washington, DC: National Science Foundation.

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.

Boston, M. D. (2006). *Developing secondary mathematics teachers' knowledge of and capacity of implement instructional tasks with high-level cognitive demands* (Doctoral dissertation, University of Pittsburgh). *Dissertation Abstracts International* (Publication No. AAT 3223943).

Boston, M. D., & Smith, M. S. (2008). Linking professional development to changes in secondary mathematics teachers' knowledge and instructional practices. Paper presentation at the annual meeting of the American Education Research Association, New York, NY.

Boston, M. D., & Smith, M. S. (2009). Transforming secondary mathematics teaching: Increasing the cognitive demands of instructional tasks used in teachers' classrooms. *Journal for Research in Mathematics Education, 40*, 119–156.

Boston, M. D., & Smith, M. S. (2011). *A 'task-centric approach' to professional development: Enhancing and sustaining mathematics teachers' ability to implement cognitively challenging mathematical tasks.* ZDM: International Journal of Mathematics Teacher Education. doi:10.1007/s11858-011-0353-2.

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher, 38*, 181–199.

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.

Fullan, M. (2009). Large-scale reform comes of age. *Journal of Educational Change, 10*, 101–113.

Grouws, D. A., Smith, M. S., & Sztajn, P. (2004). The preparation and teaching practices of United States mathematics teachers: Grades 4 and 8. In P. Kloosterman, F. Lester, & C. Brown (Eds.), *Results and Interpretations of the 1990 through 2000 Mathematics Assessments of the National Assessment of Educational Progress*. Reston, VA: National Council of Teachers of Mathematics.

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*, 524–549.

Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education, 39*, 372–400.

Hughes, E. K., Smith, M. S., Boston, M. D., & Hogel, M. (2008). Case stories: Supporting teacher reflection and collaboration on the implementation of cognitively challenging mathematical tasks. In F. Arbaugh & P. M. Taylor (Eds.), *Inquiry into mathematics teacher education.* Association of Mathematics Teacher Educators (AMTE) Monograph Series, Volume 5.

Kazemi, E., & Franke, M. L. (2004). Teacher learning in mathematics: Using student work to promote collective inquiry. *Journal of Mathematics Teacher Education, 7*, 203–235.

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*, 248–274.

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.

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

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

Scher, L., & O'Reilly, F. (2009). Professional development for K-12 math and science teachers: What do we really know? *Journal of Research on Educational Effectiveness, 2*, 209–249.

Sherin, M. G., & van Es, E. (2005). Using video to support teachers' ability to notice classroom interactions. *Journal of Technology and Teacher Education, 13*, 475–491.

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

Simon, M. A., & Tzur, R. (2004). Explicating the role of mathematical tasks in conceptual learning: An elaboration of the hypothetical learning trajectory. *Mathematical Thinking and Learning, 6*, 91–104.

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

Smith, M. S., Bill, V., & Hughes, E. K. (2008). Thinking through a lesson protocol: A key for successfully implementing high-level tasks. *Mathematics Teaching in the Middle School, 14*, 132–138.

Smith, M. S., Silver, E. A., Stein, M. K., Henningsen, M., Boston, M., & Hughes, E. (2005). *Improving instruction in algebra: Using cases to transform mathematics teaching and learning* (Vol. 2). New York: Teacher's College Press.

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 & 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*, 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*, 455–488.

Stein, M. K., Remillard, J., & Smith, M. S. (2007). How curriculum influences student learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 319–370). Greenwich, CT: Information Age Publishing.

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.

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

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

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

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.

USDE. (1999). *Exemplary and promising mathematics programs: Expert panel report*. Retrieved from: 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.

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.