The Mathematics Enthusiast

Volume 13
Number 1 *Numbers 1 & 2*

Article 1

2-2016

TME Volume 13, Numbers 1 and 2

Follow this and additional works at: <https://scholarworks.umt.edu/tme>

Part of the [Mathematics Commons](#)

Let us know how access to this document benefits you.

Recommended Citation

(2016) "TME Volume 13, Numbers 1 and 2," *The Mathematics Enthusiast*. Vol. 13 : No. 1 , Article 1.

DOI: <https://doi.org/10.54870/1551-3440.1361>

Available at: <https://scholarworks.umt.edu/tme/vol13/iss1/1>

This Full Volume is brought to you for free and open access by ScholarWorks at University of Montana. It has been accepted for inclusion in The Mathematics Enthusiast by an authorized editor of ScholarWorks at University of Montana. For more information, please contact scholarworks@mso.umt.edu.

The Mathematics Enthusiast

Volume 13, Numbers 1-2

Special issue: Mathematical Knowledge for Teaching: Developing Measures and Measuring Development

Guest Edited by Reidar Mosvold and Mark Hoover

0. Guest Editorial: Mathematical Knowledge for Teaching: Developing Measures and Measuring Development

Reidar Mosvold (Norway) & Mark Hoover (USA)

pp. 1–2

1. Making Progress on Mathematical Knowledge for Teaching

Mark Hoover (USA), Reidar Mosvold (Norway), Deborah L. Ball (USA), & Yvonne Lai (USA)

pp. 3–34

2. What Does it Take to Develop Assessments of Mathematical Knowledge for Teaching?: Unpacking the Mathematical Work of Teaching

Sarah Kate Selling (USA), Nicole Garcia (USA), & Deborah L. Ball (USA)

pp. 35–51

3. Assessing Mathematical Knowledge for Teaching: The Role of Teaching Context

Geoffrey Phelps (USA) & Heather Howell (USA)

pp. 52–70

4. Interview Prompts to Uncover Mathematical Knowledge for Teaching: Focus on Providing Written Feedback

Yeon Kim (South Korea)

pp. 71–92

5. Why Defining the Construct Matters: An Examination of Teacher Knowledge Using Different Lenses on One Assessment

Chandra H. Orrill (USA) & Allan S. Cohen (USA)

pp. 93–110

6. Knowledge for Equitable Mathematics Teaching: The Case of Latino ELLs in U.S. Schools

Aaron T. Wilson (USA)

pp. 111–129

7. In-service Teachers' Reasoning about Scenarios of Teaching Mathematics to English Language Learners

Sultan Turkan (USA)

pp. 130–148

8. Teachers and their Educators – Views on Contents and their Development Needs in Mathematics Teacher Education

Mika Koponen (Finland), Mervi A. Asikainen (Finland), Antti Viholainen (Finland), & Pekka E. Hirvonen (Finland)

pp. 149–170

9. Use of Mathematical Tasks of Teaching and the Corresponding LMT Measures in the Malawi Context

Mercy Kazima (Malawi), Arne Jakobsen (Norway), & Dun N. Kasoka (Malawi)

pp. 171–186

Guest Editorial:

Mathematical Knowledge for Teaching: Developing Measures and Measuring Development

Reidar Mosvold (reidar.mosvold@uis.no)

University of Stavanger, Norway

Mark Hoover (mhoover@umich.edu)

University of Michigan, USA

In recent years, researchers' interest in mathematical knowledge that is specific to teaching has escalated. Its increasing importance has also influenced the content of this journal. Less than two years ago, a special issue that focused on mathematical knowledge for teaching was published in *The Mathematics Enthusiast* (volume 11, no. 2). The title of this issue was: "The Mathematical Content Knowledge of Elementary Prospective Teachers". All of the articles in that special issue were written in connection with a review of 112 studies from 1978 to 2012 on the knowledge of prospective elementary teachers in the following five content areas: numbers and operations, fractions, decimals, geometry and measurement, and algebra. In their final article of that special issue, Thanheiser, Browning, Edson, Lo, Whitacre, Olanoff and Morton (2014) concluded that the number of articles on prospective teachers' mathematical content knowledge had increased over the years. They also found that most of the studies focused on "static studies of knowledge" rather than on how this knowledge developed. Fraction content knowledge was the area with most of the studies, and — across all five content areas — the studies showed that prospective teachers seemed to rely on procedural understanding, and most of the literature focused on identifying deficits in their understanding. In their conclusions, Thanheiser et al. suggested that more research should focus on characterizing prospective teachers' content knowledge and investigate how this knowledge develops.

Our aim is that the articles in the present special issue, at least to a certain extent, contribute to moving the field forward in the directions pointed out by Thanheiser et al. (2014). The focus of the articles in this special issue is on mathematical knowledge for teaching more broadly, and not only on prospective elementary teachers' knowledge. As a whole, the articles in this special issue build upon an assumption that strong instruments are of vital importance to measure — and thus to characterize the nature and development of — mathematical knowledge for teaching. Although all of the studies in this special issue draw upon the practice-based theory of mathematical knowledge for teaching that has been developed by researchers at the University of Michigan (e.g., Ball, Thames, & Phelps, 2008), the lessons learned from these studies may be relevant to the work of researchers who adhere to other views about the mathematical knowledge that is distinctly related to the teaching of mathematics.

The articles in this special issue differ in type as well as in scope, ranging across theoretical and conceptual studies, studies that focus on instruments and instrument development, and more standard empirical studies. All of them contribute either to broader efforts to develop measures of mathematical knowledge for teaching or to reflections on research that has been done and new directions important for advancing. Thus, the title of this special issue, where the first use of the word "measure" refers to the literal assigning of numbers for the purpose of comparison and the second use, "measuring

development,” refers to a figurative perusing and appraising of headway in the field. Although authors of most of the articles have some connection with the research group at the University of Michigan, the authors represent different universities and institutions from five different countries: USA, Norway, Finland, South Korea and Malawi. Many are by junior faculty and may suggest a shifting interest and focus of study taking place in the field. It is our hope that this special issue will be of interest to all researchers in the field of mathematics education, and especially to those who are involved in research on mathematical knowledge for teaching. We believe that the articles in the special issue provide a relevant snapshot of where this field of research is today, and we also hope that they provide relevant suggestions for the further development of the field.

References

- Ball, D. L., Thames, M. H. & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, 59(5), 389–407.
- Thanheiser, E., Browning, C., Edson, A. J., Lo, J.-J., Whitacre, I., Olanoff, D., & Morton, C. (2014). Prospective Elementary Mathematics Teacher Content Knowledge: What Do We Know, What Do We Not Know, and Where Do We Go? *The Mathematics Enthusiast*, 11(2), 433–448.

Making Progress on Mathematical Knowledge for Teaching

Mark Hoover

University of Michigan, USA

Reidar Mosvold

University of Stavanger, Norway

Deborah Loewenberg Ball

University of Michigan, USA

Yvonne Lai

University of Nebraska-Lincoln, USA

Abstract: Although the field lacks a theoretically grounded, well-defined, and shared conception of mathematical knowledge required for teaching, there appears to be broad agreement that a specialized body of knowledge is vital to improvement. Further, such a construct serves as the foundation for different kinds of studies with different agendas. This article reviews what is known and needs to be known to advance research on mathematical knowledge for teaching. It argues for three priorities: (i) finding common ground for engaging in complementary studies that together advance the field; (ii) innovating and reflecting on method; and (iii) addressing the relationship of such knowledge to mathematical fluency in teaching and to issues of equity and diversity in teaching. It concludes by situating the articles in this special issue within this emerging picture.

Keywords: mathematical knowledge for teaching, MKT, specialized knowledge, pedagogical content knowledge, PCK, mathematics teacher education, method, mathematical fluency, equity, diversity.

Introduction

A century ago, a central focus of teacher education in the United States was on developing a thorough understanding of subject matter, but the mid-twentieth century witnessed a steady shift to an emphasis on pedagogy generalized to be largely independent of subject matter. By the 1980s, an absence of content focus was so prevalent that Shulman (1986) referred to this as a “missing paradigm” in teacher education. A similar tendency can be seen in other countries. For example, a few decades ago, it was possible to become qualified for teaching mathematics in grades 1–9 in Norway with no more mathematics than a short course in didactics. A widespread

assumption seemed to be that prospective teachers already knew the content they needed, from their experiences as students, and they only required directions in how to teach this content. Shulman's call for increased attention to subject matter reoriented research and practice. However, the connection between the formal education of mathematics teachers and the content understanding important for their work is not straightforward. Teachers' formal mathematics education is not highly correlated with their students' achievement (Begle, 1979) or with the depth of understanding they seem to have of the mathematical issues that arise in teaching (Ma, 1999).

One of Shulman's (1986) most important contributions was the suggestion that the work of teaching requires professional knowledge that is distinctive for the teaching profession. He proposed different categories of professional knowledge for teaching. One of these categories was distinctive content knowledge, which Shulman described as including a deep knowledge of the structures of the subject (e.g., Schwab, 1978), beyond procedural and factual knowledge. Another category of knowledge was what Shulman termed "pedagogical content knowledge," which is aspects of the content most germane to its teaching (1986, p. 9). The idea about an amalgam of subject matter knowledge and pedagogical knowledge has continued to appeal to researchers working in different subject areas, and Shulman's foundational publications are among the most cited references in the field of education. (Google Scholar identifies over 13000 publications that cite his 1986 article.)

In the last two decades, researchers and mathematics educators have increasingly emphasized the significance of mathematical knowledge that is teaching-specific. Such knowledge is seen as different from the mathematics typically taught in most collegiate mathematics courses and from the mathematics needed by professionals other than teachers. Although it includes knowing the mathematics taught to students, the kind of understanding of the material needed by teachers is different than that needed by the students. Even though the literature suggests a general consensus that mathematics teaching requires special kinds of mathematical knowledge, agreement is lacking about definitions, language, and basic concepts. Many scholars draw on Shulman's notion of pedagogical content knowledge (or PCK) and view this knowledge as being either a kind of "combined" knowledge or a kind of "transformed" knowledge. Grounded in Shulman's proposals, the phrases "for teaching" and "practice-based" have been emphasized to indicate the relationship of the knowledge to specific work of teaching (e.g., Ball, Thames, & Phelps, 2008). For this article, we adopt these phrases but maintain an ecumenical view of a more extended literature.

With growing interest in ideas about specialized professional content knowledge, the early 2000s saw a spate of large-scale efforts to develop measures of such knowledge and the use of such a construct as the basis for a wide range of research studies, such as evaluating professional development (e.g., Bell, Wilson, Higgins, & McCoach, 2010), examining the impact of structural differences on the mathematical education of teachers (e.g., Kleickmann et al., 2013), arguing for policies and programs (e.g., Hill, 2011), and investigating the role of professional content knowledge on mathematics teaching practice (e.g., Speer & Wagner, 2009). Instruments for measuring such knowledge represent a crucial tool for making meaningful progress in a field. They operationalize emerging thinking, invite scrutiny, and support the investigation of underlying models.

This special issue on developing measures and measuring development of mathematical knowledge for teaching continues this focus on instruments, along with a concomitant regard for broader purposes and potential ways to advance the field. In an effort to situate this special issue, this introductory article provides some selected highlights from the field — focusing on what is being studied, how, and to what ends. To accomplish this, we draw on both a detailed review of articles sampled from 2006 to 2013 and our wider reading of the literature. We then nominate some key areas for making progress on research and development of the specialized mathematical knowledge teachers need and we use this framing to characterize the agendas and contributions of the collection of articles assembled in this special issue. The article consists of three major sections.

1. Lessons from Empirical Research
2. Next Steps for the Development of Mathematical Knowledge for Teaching
3. Articles that Develop Measures and Measure Development

The first describes a review we conducted and discusses three broad arenas of work suggested by this review. The second discusses three proposals for future research. The last briefly situates the articles in this issue within the lessons and directions discussed.

Lessons from Empirical Research

In our reading of empirical literature concerned with the distinctive mathematical knowledge requirements for teaching, several broad strands of research stand out. We begin by describing a formal review we conducted of empirical research that began appearing in about 2006, in the wake of a number of conceptual proposals (beyond PCK), and that began using these proposals as a conceptual basis for empirical study. This review informs our overall reading of the field. Combining this review with our wider reading in the field, we then identify and discuss three major arenas of work.

Reviewing the literature. In the course of other research we were conducting, we reviewed international empirical literature published in peer-reviewed journals in English between 2006 and 2013.¹ Wanting to survey the topic across theoretical perspectives, we developed and tested inclusive search terms:

- Mathematics
 - math* (the asterisk is a placeholder for derived terms)
- Content knowledge
 - know* AND (content OR special* OR pedagog* OR didact* OR math* OR teach* OR professional OR disciplin* OR domain) OR “math for teaching” OR “mathematics for teaching” OR “math-for-teaching” OR “mathematics-for-teaching

¹ This more formal review, which we use to inform our wider reading of the literature, was funded by the National Science Foundation under grant DRL-1008317 and conducted in collaboration with Arne Jakobsen, Yeon Kim, Minsung Kwon, Lindsey Mann, and Rohen Shah, who we wish to thank for their assistance with searching, conceptualizing codes, coding, and analysis. The opinions reported here are those of the authors and do not necessarily reflect the views of the National Science Foundation or our colleagues.

- Teaching
 - teaching OR pedagog* OR didact* OR instruction*

These search terms initially yielded over 3000 articles from the following six databases:

- PsycInfo
- Eric
- Francis
- ZentralBlatt
- Web of Science
- Dissertation Abstracts

Broadened search terms, additional databases, and inclusion of earlier publication years yielded none to negligible additional articles.

Based on a reading of abstracts, 349 articles were identified as potential empirical articles (as characterized by the American Educational Research Association, 2006) in which some concept of distinctive mathematics needed for teaching was used as a conceptual tool to formulate research questions or structure analysis. Our goal was not to reach high standards of reliability, but rather to use a systematic process to collect a corpus of relevant studies representing the literature from this period. In coding the articles, we sought to be descriptive rather than evaluative and iteratively worked between an inductive examination of a sample of articles and initial conceptualizations of empirical research combined with a basic model of educational change. After reading full articles, 190 of the 349 remained in the final set. A set of core codes were developed for the following categories:

1. Genre of the study
2. Research problem used to motivate the study
3. Variables used
4. Whether or not and how causality was addressed
5. Findings

Additional codes included sample size, instruments used for measuring mathematical knowledge for teaching, school level or setting, professional experience of the teachers, geographic region, and mathematical area addressed. Each article was read and coded by two team members, with a decision as to whether it satisfied our inclusion criteria, and if so, codes were reconciled. (For a more detailed description of the methods used, see Kim, Mosvold, and Hoover (2015).)

In table 1, we present some patterns that emerged from some of the additional, descriptive codes.

Table 1. *Selected descriptive codes for sample size, instrument used, level of schooling, and geographic context.*

<i>Categories and codes</i>	<i>Number of papers</i>
Sample size	
Small scale (<10)	60
Medium 1 (10–29)	51
Medium 2 (30–70)	34
Large scale (>70)	43
None	2
Instrument	
COACTIV	4
CVA	3
DTAMS	3
LMT (including adaptations)	31
TEDS-M	2
Non-standardized	56
None	91
Level of teachers	
Primary (K–8)	81
Middle (5–9)	45
Secondary (7–13)	41
Tertiary	3
Across levels	20
Regions	
Africa	7
Asia	27
Europe	22
Latin America	3
North America	112
Oceania	15
Across regions	4

We observe that many studies are small-scale, and a large number of the studies apply non-standardized instruments or no instruments. In the studies where standardized instruments were used to measure teachers' knowledge, the instruments developed to measure mathematical knowledge for teaching in the Learning Mathematics for Teaching (LMT) project were most common. An abundance of studies focuses on primary teachers, and most studies were carried out in North America.

Table 2 provides the fourteen categories developed for coding the research problem. We have grouped these into three domains and use these groups to discuss the literature in the following sections.

Table 2. *Research problems addressed.*

<i>Problems</i>	<i>Number of papers</i>	<i>%</i>
Nature and composition of SM	55	28.9
What is SM?	34	
What relationships exist among aspects of SM or with other variables?	21	
Improvement of SM	81	42.6
What professional development improves teachers' SM?	28	
What teacher education improves teachers' SM?	28	
What curriculum/tasks improve teachers' SM?	10	
What teaching practice improves teachers' SM?	0	
How SM develops?	15	
How to scale up the teaching and learning of SM?	0	
Contribution of SM	33	17.4
Does SM contribute to teaching practice?	6	
What does SM contribute to teaching practice?	12	
Does SM contribute to student learning?	15	
What does SM contribute to student learning?	0	
Other		
What SM do teachers know?	21	11.1
How policy influences teachers' SM?	0	0
<i>Total</i>	<i>190</i>	<i>100</i>

In order to make table 2 more readable, we use the abbreviation SM to signify any of the variety of ways in which mathematical knowledge for teaching might be conceptualized and named. The intention is not to introduce yet another term or acronym

for such knowledge. In this article, we have adopted more generic language to express an inclusive notion of such knowledge and we avoid the use of any specific acronym label.

For the purpose of this introductory article, we used patterns evident in the review described above to inform our extended reading of the field. Together, these efforts led us to identify three broad themes. First, a number of studies investigate the nature and composition of teacher content knowledge. Given that foundational research into teaching-specific mathematical knowledge pointed to its elusiveness and complexity, it is not surprising that scholars continue to investigate what it is — its components, measurement, features, and related constructs. A second group of studies, which constitutes the majority of published articles, investigates approaches to increasing teacher knowledge, in both the context of pre-service teacher education and the professional education of practicing teachers. A third group of studies, fewer in number, investigates effects of teachers' knowledge on both teaching and student learning. In the following sections, we use these three broad themes to organize our comments on selected highlights from the literature. Following these, we provide suggestions about possible next steps for further research in this field.

Nature and composition of mathematical knowledge for teaching. Current studies continue to probe ideas about the nature and composition of teaching-specific knowledge of mathematics. Some studies consider the construct in broad terms. They may identify or elaborate aspects or frameworks, characterize or critique the construct, compare different representations or sub-domains, or compare such knowledge with other kinds of mathematical knowledge. Others examine a constrained area of knowledge: some in relation to specific mathematical topics; some in relation to specific practices of teaching, or at specific levels (such as interpreting and responding to student thinking, curriculum use, or proving in high school geometry); and some in relation to specific qualities (such as connectedness). However, these studies do not build on each other in obvious ways and clear lessons are hard to identify. The one avenue of work that represents progress for the field is the development of instruments, and we focus our discussion there.

Instruments provide a crucial tool for investigating the nature and composition of mathematical knowledge needed for teaching. They serve to operationalize ideas about mathematical knowledge for teaching and test assumed models of the role it plays. They are used to investigate the teaching and learning of such knowledge, relationships with other variables, and other questions important for practice and policy. On the one hand, rigorous instrument development is expensive relative to budgets available for most studies and many instruments are used in a single study and limited in the extent to which they meet psychometric standards and establish validity. On the other hand, several larger efforts have invested in building instruments for large-scale studies and wider use in the field. The Learning Mathematics for Teaching (LMT) instruments for practicing elementary and middle school teachers (Hill, Schilling, & Ball, 2004) include nearly 1000 items on over a dozen different instruments and have been used in numerous program evaluations and studies of relationships and effects. They have been extensively validated (Schilling & Hill, 2007) and adapted internationally (Blömeke & Delaney, 2012). The Diagnostic Teacher Assessment in Mathematics and Science (DTAMS)

instruments for practicing middle school teachers (Saderholm, Ronau, Brown, & Collins, 2010) include 24 forms in four content areas, have been administered and rigorously analyzed with a sample of several thousand teachers, and are currently being expanded. The Teacher Education and Development Study in Mathematics (TEDS-M) instruments for pre-service primary and lower secondary teachers (Tatto et al., 2008; Senk et al. 2012) include over 100 items and were originally administered to 23,000 pre-service teachers in 17 countries.

These instruments represent an important contribution to the field. Extensive cross-professional-community review and the building of agreed-on formulations of important content knowledge have played a major role in the development of these measures. The synthesis of ideas and the integration of expertise from multiple professional communities have helped to clarify and improve ideas about mathematical knowledge for teaching. In addition, the availability of common instruments has enabled meaningful comparison and interpretation across programs, countries, and studies in ways that contribute to the maturity of research on mathematical knowledge for teaching.

Several other efforts have developed instruments with less focus on broad consensus or widespread use. The COACTIV instrument for practicing secondary teachers (Kunter, Klusmann, Baumert, Voss, & Hacfeld, 2013) produced items of a genre similar to those described above and used these to investigate relationships to other variables and to understand issues of practice and policy related to the mathematical education of teachers. Some instruments have been developed to focus on mathematical knowledge related to a specific topic, such as fractions (Izsak, Jacobson, de Araujo, & Orrill, 2012), geometry (Herbst & Kosko, 2012), algebra (McCrory, Floden, Ferrini-Mundy, Reckase, & Senk, 2012), and continuous variation and covariation (Thompson, 2015). Others have focused on specific aspects of teaching, and the mathematical knowledge required in these specific teaching practices, such as choosing examples (Chick, 2009; Zodik & Zaslavsky, 2008) and scaffolding whole-class discussions to address mathematical goals (Speer & Wagner, 2009). Many instruments have been developed in relation to specific lines of research and often in response to perceived issues with more established instruments. A number of researchers are concerned about a potentially narrow interpretation of knowledge as declarative or about a possible discrepancy between knowledge and knowledge use. These concerns have led some scholars to explore different conceptualizations of the mathematics teachers need and to look for alternative formats for measuring it (e.g., Kersting, Givvin, Thompson, Santagata, & Stigler, 2012; McCray & Chen, 2012; Thompson, 2015).

Although the development of instruments is an important step toward building a robust conception of teaching-specific knowledge of mathematics, these efforts also reveal a lack of shared language and meaning of foundational concepts. Differences in meaning for the construct PCK have been noted in the past (Ball, Thames, & Phelps, 2008; Depaepe, Verschaffel, & Kelchtermans, 2013; Graeber & Tirosh, 2008). These differences persist, yet they are often overlooked with regard to instruments.

For example, Kaarstein (2014) examined whether the LMT, TEDS-M, and COACTIV instruments, each referencing Shulman and stating that the respective instrument measures PCK, measure the same thing. To study this issue, she constructed a taxonomy of the different levels of categories in Shulman's initial framework as well as

the frameworks that were used to develop the three instruments. She then selected three items — one from each instrument — and categorized them according to each of the three frameworks. Her main argument is that content knowledge and pedagogical content knowledge are supposed to be distinct categories, and therefore three projects that use the same basic categories should categorize items in the same way. However, from her analysis the items would be placed in different basic categories using the criteria reported by the projects. As an example, an item that was categorized as a specialized content knowledge item (measuring content knowledge) in the LMT project would probably have been categorized as a PCK item in TEDS-M and COACTIV. Kaarstein's argument does not necessarily threaten the validity of the measures from each of the three projects, but her observation deserves further attention.

Similarly, a study by Copur-Gencturk and Lubienski (2013) echoes this concern. In order to investigate growth in pre-service teacher knowledge, they used two different instruments: LMT and DTAMS. When comparing groups of teachers who had participated in different kinds of courses, they concluded that the LMT and DTAMS instruments measure aspects of mathematical knowledge for teaching that are substantially different. Teachers who participated in a hybrid mathematics content/methods course had the most significant increase in their LMT score, and this score remained stable although they took an additional content course. Teachers' DTAMS score also increased during the hybrid course; during the content knowledge course, only the content knowledge part of their DTAMS increased. This study thus supports the idea that there is specialized mathematical content knowledge not influenced by general mathematics content courses. That different instruments measure different aspects of knowledge is not necessarily surprising, but it is worrying if instruments ostensibly designed to capture the same construct in fact measure significantly different facets of that knowledge, with little clarity about these differences.

The concerns raised by Kaarstein (2014) and Copur-Gencturk and Lubienski (2013) suggests that the limited specification of the construct and the different ways of operationalizing it makes it difficult to interpret results. This limits the extent to which results from these instruments, taken individually or together, can inform the conceptualization of mathematical knowledge for teaching or practical decisions needed to design learning opportunities.

Developing teachers' mathematical knowledge for teaching. With a growing sense of the mathematics important for improving teaching and learning, practitioners have turned their attention to increasing teachers' knowledge of professionally relevant mathematics and scholarly work has followed suit. A large number of studies make it clear that the design and evaluation of teacher education and professional development programs in developing teachers' mathematical knowledge for teaching are top priorities. From several decades of research, we propose what we see as a few related emerging lessons:

- Teaching teachers additional standard disciplinary mathematics beyond a basic threshold does not increase their knowledge in ways that impact teaching and learning.

- Providing teachers with opportunities to learn mathematics that is intertwined with teaching increases their mathematical knowledge for teaching.
- The focus of the content, tasks, and pedagogy for teaching such knowledge requires thoughtful attention to ways of maintaining a coordination of content and teaching without slipping exclusively into one domain or the other.

These lessons are rooted in early efforts to document effects of teachers' mathematical knowledge on student learning and are reinforced by current research on the design and implementation of teacher education and professional development. We begin by briefly reflecting on that early work and then tracing these lessons into current research.

Much of the impetus for the surge in research on teaching-specific knowledge began with reviews of several decades of large-scale research that found surprisingly little to no effect of teachers' mathematical knowledge on their students' learning (Ball, Lubienski, & Mewborn, 2001). The studies reviewed were often conducted with large datasets but very coarse measures. Taking Shulman's (1986) suggestion that the content knowledge needed by teachers was characteristically different from that needed by other professionals, researchers began to look more closely at the measures used in those studies and at the findings. The clearest finding that emerged was that methods courses consistently showed positive effects while content courses did not (e.g., Begle, 1979; Ferguson & Womack, 1993; Guyton & Farokhi, 1987; Monk, 1994). The second was that positive effects were more likely when the content taught to teachers was more closely related to the content they subsequently taught. For instance, several scholars found effects when using student exams to measure teachers' knowledge (Harbison & Hanushek, 1992; Mullens, Murnane, & Willett, 1996). Reinforcing these results, Monk (1994) found that coursework in calculus influenced the achievement of secondary teachers' students in algebra classes, but not in their geometry classes. In general, when the mathematics taught or measured is meaningfully connected to classroom materials or interactions, it is modestly associated with improved teaching and learning.

For some practitioners and policy-makers, the implication of these empirical studies, combined with logical arguments for teaching-specific professional knowledge, has been enough to lead to prioritizing mathematical knowledge for teaching in the mathematical education of teachers. Nevertheless, many policies continue to press for increases in the number of mathematics courses required of teachers, regardless of their connection to teaching, despite abundant evidence that such policies are unlikely to improve teaching and learning (e.g., Youngs & Qian, 2013). Such policies have probably been less the result of lingering doubt about empirical results and more the result of overextending the notion that knowing content well is key to good teaching, even in the face of disconfirming evidence. Of course, a certain threshold level of knowledge of the subject is essential, but preparing teachers by requiring mathematics courses that are not directly connected to the content being taught or to the work involved in teaching that content is misguided.

More recent studies continue to reinforce these established lessons. One recent line of inquiry is the investigation of features of innovative, well-received professional development programs. To us, the most compelling result emerging from these studies is that professional development requires designing pedagogically relevant movement

between mathematical and pedagogical concern both to motivate teachers' investment in mathematical issues and to keep the mathematical attention on mathematics that matters for the work of teaching. To elaborate, we offer several examples that contribute to this claim.

With deep regard for the limited effects of decades of substantial national investment in professional development, several research groups have organically developed approaches informed by thoughtful reflection and attention to disciplined observation of teachers' engagement with and actual uptake of ideas and practices. One important insight emerging from these decades-long investments is that cycling through mathematical considerations, pedagogical considerations, and reflective enactment is vital to the design of professional development. For instance, Silver, Clark, Ghouseini, Charalambous, and Sealy (2007) set out to provide evidence for whether and how teachers might enhance their mathematical knowledge for teaching through monthly practice-based professional development workshops designed to cycle from activities of doing mathematics, to examining case-based pedagogical and student-related issues, to planning, teaching and debriefing lessons collaboratively (all related to a common mathematics task or set of tasks). Examining the interactions of one teacher, they document ways these activities provided opportunities for teachers to build connections among mathematical ideas and to consider these ideas in relation to student thinking and teaching. They do not measure teacher learning. Nor do they disentangle effects of what they refer to as a professional-learning-task cycle from a number of other important features of their professional development program. However, they document dynamics in which the teacher, from an initial experience solving a nontrivial mathematics problem, supported by mathematically sensitive facilitation, successively engages in mathematical issues and pedagogical issues in ways that visibly build connections among mathematical ideas, pedagogical practice, and growing mathematical knowledge for teaching. In addition, they argue that their cyclic design increased teachers' motivation for learning mathematics, both in the workshops and in their daily practice.

Through successive opportunities to consider mathematical ideas in relation to the activities of classroom practice, our participants came to see their pedagogical work as permeated by mathematical considerations. (p. 276)

Similarly, in working to close the gap between a reform vision and the actual practice of mathematics teaching and learning, Koellner et al. (2007) implemented a model of professional development designed to help teachers deepen their mathematical knowledge for teaching through a cycle of solving a mathematics problem, teaching the problem, and analyzing first teacher questioning and then student thinking in videos of their teaching. In order to understand the learning opportunities afforded by what they refer to as a problem-solving-cycle design, they analyzed artifacts from two years of a series of monthly, full-day workshops with ten middle school mathematics teachers, including workshop videos and interviews with facilitators. The researchers used the knowledge domains identified in Ball, Thames, and Phelps (2008) to analyze several teacher interactions. They found that different learning opportunities were afforded by different activities: specialized content knowledge was developed by comparing, reasoning, and making connections between the various solution strategies; knowledge of

content and teaching was developed by analyzing teacher questioning in the video clips from the teachers' lessons; and knowledge of content and students was developed by analyzing students' solution methods (interpreting them and considering their implications for instruction). More importantly, the researchers found that reflecting on and discussing the nature of student thinking and teacher questioning of students evident in videos of their own teaching led teachers to extend their mathematical knowledge for teaching as they re-engaged with the mathematics problem and reconsidered how they might teach the problem in light of their new regard for how students might approach the problem. Throughout the analysis, the authors found that specialized content knowledge interacts with pedagogical content knowledge in interpreting student thinking and planning lessons. The authors argue that the workshops developed teachers' mathematical knowledge for teaching by supporting teachers' current knowledge, while gradually challenging them to gain new understanding for the purpose of their work as teachers.

The lessons from studies such as these are subtle. The movement between mathematical study and pedagogical practice is central, but attention needs to be given to dynamics regarding teachers' motivation, the timing of different activities, and specific mathematical opportunities arising from specific pedagogical activities. In reading these reports, one gets the sense that really smart enactment of the professional development was key to success and that replicating effects might be challenging. From this work, it would seem important to discern the essential design features and elaborate the necessary character of facilitation.

One effort along these lines is a study by Elliott, Kazemi, Lesseig, Mumme, and Kelley-Petersen (2009). In the context of supporting facilitators' enactment of mathematically focused professional education, they analyzed facilitators' learning and the use of two frameworks provided as conceptual tools: (i) sociomathematical norms for cultivating mathematically productive discussion in professional development, adapted from Yackel and Cobb (1996) and (ii) practices for orchestrating productive mathematical discussions, adapted from Stein, Engle, Smith, and Hughes (2008). In their study, Elliot and colleagues collected extensive documentation and analyzed the learning of 5 of the 36 facilitators trained at two sites in 6 two-day seminars across an academic year. They found that although facilitators responded positively to the frameworks, they experienced tensions in using the frameworks to ask questions about colleagues' mathematical thinking and they struggled with the fact that teachers positioned themselves and others as better or worse in mathematics. These dynamics got in the way of productive mathematical discussions and frustrated facilitators. The analysis revealed that one way to mitigate these tensions was by helping facilitators to identify mathematical ideas that teachers would readily see as worth developing. This led the researchers to see a need for developing more nuanced and detailed purposes for doing mathematics in professional development in ways that teachers would see as relevant to their work.

This then led the researchers to realize that they needed a way to focus the purpose and work of professional development on connections between mathematics and the work of teaching. To accomplish this, they added a third framework to their design. The authors argue that the mathematical-knowledge-for-teaching framework engaged

facilitators in understanding the ways in which specialized content knowledge (SCK) connects mathematics to teaching and that the framework provided a meaningful articulation of the purpose of the professional development and a helpful focus for the mathematical tasks and discussions that took place.

By understanding how a SCK-oriented purpose for PD is tied to classroom teaching and being able to articulate that understanding to teachers in accessible ways, leaders will be able to begin to address the pressure they felt to assure relevance in their PD. (Elliott et al., 2009, p. 376)

Again, the dynamics between mathematics and the motivation and use of that mathematics is key to effective teacher learning of professionally relevant mathematics.

The field is also beginning to see evidence that these insights have measurable yield. For instance, Bell, Wilson, Higgins, and McCoach (2010) argue that it is the practice-based character of the nationally disseminated Developing Mathematical Ideas (DMI) mathematics professional development program that best explains participating teachers' learning of mathematical knowledge for teaching. The researchers examined pre and post teacher content knowledge for 308 treatment and comparison teachers across 10 well-established sites. They found significantly larger gains for treatment teachers' scores and that these gains were related to breadth of opportunity to learn provided by facilitators. Methodically considering a number of alternative explanations for treatment teachers' improvement, the researchers emphasize the classroom-practice feature of the professional development, where teachers move back and forth between seminars and their own classrooms, receiving written feedback from regularly observing facilitators. Referring to Ball and Cohen's (1999) argument that teacher learning needs to be embedded in practice, they point out that connecting to practice can leverage teacher learning in and from their daily work, greatly expanding overall capacity for teacher learning and improvement. They argue that the practice-based nature of their design contrasts with professional development that takes place apart from teachers' practice.

DMI is quite different in this regard, for it encourages teachers to take their nascent SCK, KCS, and KCT into their classrooms and try things out. Repeatedly, teachers told us of their revelations — both in seminars and in their own schools — as they drew on their growing knowledge of and enthusiasm for mathematics and teaching mathematics in their classrooms. This anecdotal evidence aligns with results from S. Cohen's (2004) yearlong study of changes in teachers' thinking and practices over the course of their participation in DMI seminars. (Bell et al., 2010, p. 505)

These different studies compellingly add to the arguments that teachers need mathematical knowledge that is connected to the work they do and that situating the learning of mathematical knowledge in teachers' practice supports the learning of mathematical knowledge for teaching. Bell et al.'s (2010) large-scale study of the effect of professional development on teacher learning corroborates the qualitative, small-scale findings of the other studies. The professional development models highlighted set teachers up to learn in and from their practice. Together, the studies discussed above point to the coordinated nature of mathematical knowledge for teaching and the ways in

which the coordination between mathematics and pedagogy is essential to teaching and learning mathematical knowledge for teaching.

Impact of mathematical knowledge for teaching. Whereas more studies have investigated the nature and composition of mathematical knowledge for teaching and developing teachers' knowledge, fewer studies have investigated the impact such knowledge has on teaching and learning. As mentioned earlier, several studies report positive effects of mathematical knowledge for teaching on student learning. Crucial to this research has been the development of robust instruments assessing mathematical knowledge for teaching. The field has found evidence linking mathematical knowledge for teaching to student achievement using the LMT instrument (e.g., Hill, Rowan, & Ball, 2005; Rockoff, Jacob, Kane, & Staiger, 2011), the COACTIV instrument (e.g., Baumert et al., 2010; Kunter et al., 2013), and the Classroom Video Analysis (CVA) instrument (e.g., Kersting et al., 2010, Kersting et al., 2012). A fewer number of studies have investigated links between teaching practice and mathematical knowledge for teaching and/or student achievement (e.g., Hill, Kapitula, & Umland, 2011). In these studies, student learning is mostly measured by standardized test scores, and the studies vary in how they measure teaching quality. These studies indicate that, generally speaking, mathematical knowledge for teaching impacts teaching and learning.

We acknowledge the importance of studies that identify an influence of teachers' mathematical knowledge on teaching and learning, but are particularly excited about studies that unpack the dynamics of how mathematical knowledge for teaching impacts teaching and learning. In their study of 34 teachers, Hill, Umland, Litke, and Kapitula (2012) demonstrated that the connection between mathematical knowledge for teaching (measured with the LMT instrument) and the quality of instruction is complex. While weaker mathematical knowledge for teaching seemed to predict poorer quality of instruction, and stronger mathematical knowledge for teaching seemed to predict higher quality of instruction, teachers who performed in the midrange on the LMT measure varied widely in the quality of their instruction. Student achievement also varied widely for teachers with mid-range mathematical knowledge for teaching. Furthermore, Hill et al.'s (2008) study of 10 teachers found that although use of supplemental curriculum materials, teacher beliefs, and professional development are factors of potential influence, these factors might all cut both ways depending on the teachers' mathematical knowledge for teaching. These two studies underscore that simply establishing impact of knowledge on teaching is not enough to make decisions about teacher education or policy.

To frame a fuller consideration of impact, we reflect briefly on the nature of teaching and learning. Teaching mathematics involves managing instructional interactions, including everything teachers say and do together with students focused on content, where teacher knowledge is a resource for the work (Cohen, Raudenbush, & Ball, 2003). This observation suggests that in addition to general effect studies on teaching and learning, it would be helpful to know more about which specific aspects of teaching and learning are influenced by teacher content knowledge, which specific aspects of teacher content knowledge are influential, and how the influences impact interactions among teacher and students around content. In other words, we propose that Cohen et al.'s conceptualization of teacher content knowledge as a resource impacting

instructional interactions is important for framing an investigation of mathematical knowledge for teaching suited to informing the improvement of teaching and learning.

A promising direction in recent work has been initial investigation of the specific influence that mathematical knowledge for teaching has on teaching. One example of this kind is Speer and Wagner's (2009) case study of one undergraduate instructor's scaffolding of classroom discussions. Using Williams and Baxter's (1996) constructs of social and analytic scaffolding as a frame, Speer and Wagner argue that aspects of pedagogical content knowledge are important for helping students find productive ways of solving particular problems and for understanding which student contributions — correct or incorrect — are important to emphasize in a discussion. They trace ways in which particular knowledge of students' understanding aids teachers in assuring that the lesson reaches intended mathematical goals and in understanding the role of particular mathematical ideas in students' development.

In a similar vein, an exploratory study by Charalambous (2010) investigated teachers' knowledge in relation to selection and use of mathematical tasks. He investigated the teaching of two primary mathematics teachers with different levels of mathematical knowledge for teaching and found notable differences in the quality of their teaching. He used Stein and colleagues' mathematical tasks framework to examine the cognitive level of enacted tasks, and he formulated three tentative hypotheses about mechanisms of how mathematical knowledge for teaching impacts teachers' selection and use of mathematical tasks. First, he hypothesizes that strong mathematical knowledge for teaching may contribute to a use of representations that supports students in solving problems, whereas weaker mathematical knowledge for teaching may limit instruction to memorizing rules. Second, he proposes that mathematical knowledge for teaching appears to support teachers' ability to provide explanations that give meaning to mathematical procedures. Third, he proposes that teachers' mathematical knowledge for teaching may be related to their ability to follow students' thinking and responsively support development of understanding.

These two studies exemplify potential analyses of mathematical knowledge for teaching in relation to frameworks of teaching and learning. They leverage findings about teaching to probe the contributions of mathematical knowledge for teaching in ways that begin to unpack the specific role such knowledge plays. They are not the only studies to do so, but to date studies in this realm are rare. Building on these ideas, further conceptualization of distinctly mathematical tasks of teaching might provide even more focused contexts for studying mathematical knowledge for teaching as a resource for teaching. Establishing agreed-upon conceptualizations of mathematical knowledge for teaching related to well-studied components of the work of teaching and using these as a common ground for instrument development would provide a solid foundation for advancing the field.

From this brief review of recent progress on identifying, developing, and understanding the impact of mathematical knowledge for teaching, we now turn our attention to proposing directions for future work.

Next Steps for the Development of Mathematical Knowledge for Teaching

As described above, compelling examples of mathematical knowledge for teaching and evidence associating it with improved teaching and learning have sparked interest in making it a central goal in the mathematical education of teachers. However, various impediments exist. The lack of rigorous, shared definitions and the incomplete elaboration of a robust body of knowledge create problems for meaningful measures and curricula development. Underlying these challenges are competing ideas about how to conceptualize the knowledge, questions about the relationship among knowledge, knowledge use, and outcome, and the need for ways to decide claims about whether or not something constitutes professional knowledge.

We suggest three priorities for research and development of mathematical knowledge for teaching: (1) focused studies that together begin to compose a more coherent, comprehensive, and shared understanding of what it is, how it is learned, and what it does; (2) innovation and reflection on method for investigating it; and (3) studies of mathematical fluency in teaching and the nature of mathematical knowledge for equitable teaching. Below, we argue that each of these is vital to long-term progress in improving the mathematical education of teachers and the mathematics teaching and learning that depends on it.

Investigating focused issues while contributing to a larger research program.

Scores of articles in the previous decade have argued for particular ways of distinguishing and conceptualizing important knowledge, and many others have sought to establish its presence and overall impact. With a sense of the importance of mathematical knowledge for teaching, additional studies explored the teaching of such knowledge. However, on the whole, conceptual work has been exploratory, measures have been general, and studies of the mathematical education of teachers have been limited by under-specification of the body of knowledge. We suggest that the field would benefit from focused studies that build on each other in ways that begin to put in place the machinery needed to develop an overall system for educating teachers mathematically. Such a system would include clear content-knowledge standards for professional competence, comprehensive content-knowledge course and program curricula, robust exit or professional content-knowledge exams, and rationale for what is to be taught in pre-service programs and what is better addressed in early career professional development or later on. To get there, we propose collectively pursuing several focal areas of study.

First, mathematical knowledge for teaching needs to be elaborated — for specific mathematical topics and tasks of teaching, across educational levels. Some of this work is underway, but we suggest that more needs to be done in ways that research studies, taken together, define a body of professional knowledge and provide a basis for curricula, standards, and assessments. One area of need that stands out is the investigation of the mathematical knowledge demands associated with particular domains of the work of teaching, such as leading a discussion, launching students to do mathematical work, or deciding the instructional implications of particular student work. This is a particularly challenging area of study because the field lacks comprehensive, robust specifications of the work of teaching. It is also a potentially promising area of study. Where initial decompositions of teaching are available, such as for orchestrating discussions, awareness of the mathematical knowledge entailed in the teaching can position teachers

to learn both the domain of teaching and the mathematical knowledge more productively (Boerst, Sleep, Ball, & Bass, 2011; Elliott et al., 2009). Nonetheless, domains of teaching need additional parsing before they can be fully leveraged.

A second proposed area of study is determining meaningful “chunking” of mathematical knowledge for teaching and practical progressions for teaching and learning it. In considering the mathematics that students need to learn, topics are typically decomposed into a sequence of small-sized learning goals. In contrast, teachers’ mathematical knowledge for teaching is not simply a mirror image of student curriculum. Teachers need knowledge that is different in important ways from the knowledge students need to learn. Mathematical knowledge for teaching is related to student curriculum, but it is not clear what this relationship implies for how it is best organized. In contrast to the mathematics that students need to learn, the specialized mathematics that teachers need to learn appears to be constituted in ways that span blocks of the student curriculum.

For instance, a teacher who learns how to model the steps of the standard addition algorithm using base ten blocks might still need to think through modeling subtraction, but as a minor extension of what is already learned, not as a new topic, requiring a new program of instruction. The question deserves more careful examination, but our experience is that teachers who participate in professional development related to a particular strand of work on place value exhibit significantly increased mathematical knowledge for teaching more generally across whole number computation, but with little to no impact on their mathematical knowledge for teaching topics related to geometry, data analysis, or even rational number computation. This is just a conjecture, but we offer it as a way to indicate an area of study that would contribute to improved approaches to the mathematical education of teachers. How big are these chunks? What are possibilities for structuring the chunks? Which have the greatest impact for beginning teachers? Some of these questions could be investigated as part of the elaboration research described above. Our point is that beyond the important goal of identifying knowledge for specific mathematical topics and tasks of teaching, across educational levels, research on how best to organize that knowledge might usefully inform the mathematical education of teachers.

This discussion leads to a third proposed line of investigation, one that explores mathematical knowledge for teaching along a professional trajectory from before teachers enter teacher preparation, through their training and novice practice, and into their maturation as professionals. This would require navigation among questions about what teachers know, what might be learned when, what is essential to responsible practice, and what can be sensibly coordinated with growing professional expertise. For this, the field would need to know more about the mathematical knowledge for teaching that prospective teachers bring to teacher education and whether there are things that might more readily be learned in the program and others that might be more productively required before admission. The field would need to know more about mathematical knowledge for teaching that is readily acquired from experience, as well as the supports needed to do so. Researchers would need to investigate how to distinguish between the mathematical knowledge for teaching that is essential to know before assuming sole

responsibility for classroom instruction and the knowledge that can be safely left to later professional development. We suggest that such studies would contribute to developing coherence, efficiency, and responsibility in an overarching picture of the mathematical education of teachers.

Another proposed area of study would extend work that examines effects of specific mathematical knowledge on specific teaching and learning in ways that identify underlying mechanisms and informs views of when and how mathematical knowledge is used in teaching. We noted above a need for more studies that unpack relationships among mathematical knowledge for teaching, teaching practice, and student learning. Such studies might examine the nature of student learning gains resulting from specific teacher knowledge or they might investigate the mechanisms by which teachers' mathematical knowledge for teaching has an impact. They would provide a better understanding of the nature and role of mathematical knowledge in teaching, informing both its conceptualization and validating underlying assumptions about its significance.

Finally, we suggest that the field would benefit from more studies of effects at a mid-range level, above that of idiosyncratic, individual programs and courses and below that of large-scale, international studies. In their 2004 International Congress on Mathematics Education plenary, Adler, Ball, Krainer, Lin, and Novotna (2008) observed that the majority of studies in teacher education are small-scale qualitative studies conducted by educators studying the teachers with whom they are working within individual programs or courses. The TEDS-M study and the development of some of the instruments described above have supported an increase in large-scale and cross-case studies, but as Adler and her colleagues point out, the study of courses, programs, and teachers by researchers who are also the designers and educators of those programs and teachers creates both opportunities and risks. From our review, our sense is that many small studies are driven by convenience and reduced cost, but at the expense of rigorous design and skeptical stance. Mid-sized studies would be enhanced by efforts such as developing collaborative investigations across remote sites with either similar or contrasting interventions. This is consistent with arguments about research on professional development made by Borko (2004).

Next, we argue that the agenda sketched above will require explicit development of methods for conducting such research efficiently and effectively.

Innovating and reflecting on method.² We propose that a central problem for progress in the field is a lack of clearly understood and practicable methodology for the study and development of mathematical knowledge for teaching. First, many researchers, including graduate students, seem eager to conduct studies in this arena, but choices about research design and approaches to analysis are uncertain. In our review of the literature, we found that methods vary widely, are relatively idiosyncratic, and are in general weak — in some cases attempting to make causal claims from research designs poorly suited for such claims and in others providing thoughtful claims but from unclear

² Material in this section is based on work supported by the National Science Foundation under grant number 1502778. Opinions are those of the authors and do not necessarily reflect the views of the National Science Foundation.

processes and underdeveloped logical rationale. We suspect that a lack of clarity and rigor of methods, including in our own work, are a result of several factors: unresolved and underdeveloped conceptualization of the terrain; competing purposes of research (often within a single study); and uncertain grounds for making claims about whether something does or does not constitute professional knowledge. Struggles to design robust studies and to articulate methods used suggest a need for increased attention to method.

This should not be surprising. The vitality of research in areas still in early stages of theory development requires a concomitant consideration of method. Importing method from other arenas is appropriate, but regard for the theoretical foundations of the object of study and their implications for all aspects of method is also important. We propose that reflective innovation of method, grounded in emerging theory of teaching, can better account for confounding variables that are relevant to teaching and can inform the alignment among research questions, design, analysis of data, claims, and interpretations. To ground this proposal, we reflect on two approaches that have been evident in efforts to study the nature and composition of mathematical knowledge for teaching (interview studies and observational studies) and then suggest directions for potential innovations.

Early investigations of teacher content knowledge were mostly limited to correlational studies (e.g., Begle, 1979). Correlational studies remain prominent in the field (e.g., Baumert et al., 2010; Hill, Rowan, & Ball, 2005; Kersting et al., 2010), but in the 1980s and 1990s studies began using teacher interviews to investigate teacher knowledge (see Ball, Lubienski, & Mewborn, 2001). This early work was limited in two ways. First, it tended to focus on identifying deficits in teachers' mathematical knowledge instead of clarifying the mathematical knowledge requirements of teaching. Second, although some good interview prompts emerged and supported a surfeit of studies, generating additional high-quality prompts has not been easy. The strength of these early interview prompts was that they were focused, specific, and offered compelling examples of specialized mathematical knowledge that would be important for teachers to know. The weaknesses were that they focused the conversation on teachers' lack of knowledge, while providing little insight into how to rectify these lacks, and they left the difficult work of generating good prompts invisible.

Similarly, methods for observational studies have often been weakly specified and hard to use by other scholars as the basis for complementary study. For example, because of the shortcomings of teacher interviews, the research group at the University of Michigan developed a practice-based approach to the study of video records of instruction (Ball & Bass, 2003; Thames, 2009). This approach requires simultaneously conceptualizing the work of teaching together with the mathematical demands of that work. It is empirical, interdisciplinary, analytical-conceptual research that involves developing concepts and conceptual framing by parsing the phenomenon and systematically testing proposals for consistency with data and with relevant theoretical and practice-based perspectives. The approach is time-intensive and expensive, requires skillful use of distributed expertise, and is sufficiently underspecified to make broader use challenging. For instance, early on, these researchers wrote about ways in which inter-disciplinary perspectives were central to their analyses (e.g., Ball, 1999; Ball &

Bass, 2003), but this characterization, although it captures an important feature of the approach, is inadequate as a characterization of their approach and as a method for others to use. It is underspecified, relies more on experienced judgment than on independently usable criteria or techniques, and leaves key foundational issues in doubt (Thames, 2009).

Reflecting on our own use of these approaches, we offer several, somewhat ad hoc, observations.

- Teaching is purposeful work and, as such, imposes logical demands on the activity, and these logical demands play a role in warranting claims about the work of teaching and mathematical knowledge needed for teaching.
- Mathematical knowledge for teaching is professional knowledge, and central to its development and articulation is professional vetting or consensus building based on cross-community professional judgment.
- The pedagogical context provided in crafted items and prompts entails engagement in the work of teaching and in the use of mathematical knowledge and, as such, provides crafted instances for the study of specialized knowledge for teaching.
- There is an iterative process among the development of instructional tasks, assessment items, and interview prompts and our increasing capacity for eliciting and engaging mathematical knowledge for teaching.
- Analysis of mathematical knowledge as professional knowledge for teaching, whether in situ or in constrained instances, is fundamentally an empirical, conceptual-analytic, normatively informed process, not a strictly descriptive one.

We believe that the first two observations have important methodological implications that are as of yet unrealized. Key to understanding teaching and its knowledge demands is understanding its contextual rationality. In other words, meaningful study of teaching must account for the directed and contextual nature of the work. We suspect that such study will require the development and use of methods fit to the work of teaching and that this means greater reliance of underlying theory of teaching in designing studies and choosing methods of analysis. As Gherardi (2012, p. 209) succinctly summarizes in her writing about conducting practice-based studies, “Hence, empirical study of organizing as knowing-in-practice requires analysis of how, in working practices, resources are collectively activated and aligned with competence.” She argues for thoughtful consideration of how methodological approaches are positioned in relation to the nature of practice and its constant reconstitution in the context of professional work. We agree with this position and suggest that it is exactly these issues that need to be taken up in an investigation of method for studying mathematical knowledge for teaching.

The second observation in our list above raises additional concerns for the development of new methods. Mathematical knowledge for teaching is professional knowledge, in the sense that it is shared, technical knowledge determined by professional judgment (Lortie, 1975; Abbott, 1988), but it is distinctive as a body of knowledge in that it requires the coordination of mathematics with teaching, which are different areas of expertise resident in different professional communities (Thames, 2009). Thus, the study of such knowledge requires coordination across different professional communities with

different disciplinary foundations. In other words, the study of mathematical knowledge for teaching requires, or is at least enhanced by, collective work across distinct professional communities with different expertise and different professional norms and practices, and such work requires special consideration and support (Star & Griesemer, 1989).

With the call for cross-professional coordination, the study of mathematical knowledge for teaching involves much more than an assembly line model where different professional constituents inject their specific expertise into a product handed down the line. It calls for specification of processes for collectively considering whether a proposed claim of professional knowledge is warranted. It is about establishing protocols for merging and melding different expertise in the midst of improvement work that attends to overall coherence and practical merit. It requires specific ways of working together, tools for organizing the scholarly work, and boundary objects that provide meaningfully bridges among communities (Akkermann & Bakker, 2011). Each of these adds to the need for new methodology.

Innovation and reflection on method can be carried out in numerous ways. Researchers can simply attend more closely to decisions of method and explicit reporting of method. Alternatively, they can deliberatively develop, implement, and study methods. In order to provide a sense of the kind of innovation and reflection that might be done, we discuss some of the ways we have begun to explore methodological approaches for the study of mathematical knowledge for teaching.

An emerging approach we find promising is to use sites where professional deliberation about teaching are taking place as sites where we might productively research the work of teaching and its mathematical demands. In recent studies, we have designed interview protocols as a tool for generating data useful for studying the mathematical work of teaching. For instance, to investigate the work involved in providing students with written feedback, Kim (this issue) provides a strategic piece of student work and asks interviewees to provide written feedback and to explain the rationale for the feedback. Here, instead of videotaping classroom instruction and analyzing the mathematical demands of teaching, Kim analyzes those demands as they play out in a constrained slice of the mathematical work of teaching as evidenced in responding to a teaching scenario provided.

We see this approach as an instance of a more general phenomenon, one of using sites of professional deliberation about teaching as research sites for studying teaching. For example, a group of mathematics teachers and mathematics educators in a professional development setting might discuss responses to a particular pedagogical situation in ways akin to the pedagogical deliberations of a teacher engaged in teaching. Thus, this professional development event can be useful for studying professional practice. It may even have the advantage that professional action and reasoning are more explicitly expressed, yet of course, with certain caveats in place as well, such as recognizing that real-time demands of teaching are suspended. Similar opportunities can arise in other settings where pedagogical deliberations take place, such as teacher education or the development of curriculum or assessment. For instance, recent investigation of the design process for producing tasks to measure mathematical

knowledge for teaching suggests that writing and reviewing such tasks can provide insight into teaching and its mathematical demands, even to the point of serving as a site for investigating mathematical knowledge for teaching (Jacobson, Remillard, Hoover, & Aaron, in press; Herbst & Kosko, 2012).

We propose that such an approach is distinctively different from general interview techniques that have teachers reflect on their teaching. Crucial to this difference is that the prompts are designed to provide authentic pedagogical contexts with essential, yet minimal, constraints for directing targeted pedagogical work (such as a crucial instructional goal, a key excerpt from a textbook, or strategically selected student work). Good pedagogical context needs to be based on initial conceptions of key aspects of the work, and constraints need to be designed to engage initial ideas about the nature and demands of the work. Otherwise, the pedagogical context of the tasks is unlikely to engage people in authentic pedagogical work.

Our recent experience with interview prompts of this kind has convinced us of their potential for studying teaching and teacher knowledge. Several advantages are evident: constraints provided can be manipulated; different professional communities can be engaged; and bounded instances of work examined. The development of this approach would support new lines of research that specify teaching and its professional knowledge demands in ways that can better inform professional education and evaluation. They are also easy to use and require only modest time and expense.

Such innovations begin to suggest the development of a “laboratory science” approach for studying mathematical knowledge for teaching that takes advantage of the tools of constrained prompts, the generative analytic techniques of instructional analysis, and the multiple sites available for such study. By a “lab science” approach we mean direct interaction with the world of instruction or slices of instruction using tools, data collection techniques, and models and theories of teaching. Analogous to the ways in which experimental psychologists isolate phenomena under controlled conditions in a laboratory setting or biochemists manipulate protein processes at the bench, we propose that the study of specialized teacher content knowledge can isolate activities of teaching and the use of resources, examine those activities and resources in detail, and systematically manipulate constraints to better understand phenomena. This work can be done deductively, to test specific hypotheses, inductively, to discern functional relationships, or abductively, to refine current understanding. Such an investigation of method should be intimately tied to underlying foundational issues, both shaped by theoretical commitments and giving precise definition and form to underlying theory.

In conclusion, we suggest that the development of usable, practical, and defensible method, whether along the lines we have sketched here or along other lines, will be critical to carrying out the extensive agenda described earlier for building a understanding of mathematical knowledge for teaching adequate for sustainable improvement of the mathematical education of teachers. We now sketch two areas of study largely missing from the literature on knowledge distinctive for teaching mathematics and argue that both of these are essential to viable progress on building a theory and practice of mathematical knowledge for teaching.

Addressing Two Key Issues: Mathematical Knowledge for Fluent and Equitable Teaching. Although there has been substantial progress in conceptualizing and understanding the mathematical understanding needed for the practice of teaching, significant issues remain. We focus here on two aspects that seem to us to be particularly critical to progress on mathematical knowledge for teaching. One centers on the communicative demands of teaching, the other on what is involved in teaching to disrupt the historical privileging of particular forms of mathematical competence and engagement, resulting in persistent inequity in access and opportunity. We argue that both of these are key to the long-term viability of efforts to improve the mathematical education of teachers.

Teaching is inherently a communication-intensive practice. Teachers listen to their students, explain ideas, and pose questions. They read their students' written work and drawings, and provide written feedback. Throughout these communications, they use mathematics in a range of specialized ways. They must hear what their students say, even though students talk and use mathematical and everyday language in ways that reflect their emergent understanding. Similarly, they must interpret students' writing and drawings. When they talk, teachers must attune their language to students' current understanding, and yet do so in ways that are intellectually honest and do not distort mathematical ideas to which they are responsible for giving their students access.

What is involved in this sort of mathematical communication in the context of teaching? Because teaching is fast-paced and interactive, the demands are intense. Talk and listening cannot be fully scripted or anticipated. A special kind of mathematical fluency is required, tuned to the work of teaching. Asking a question in the moment; explaining in response to a student's puzzlement; listening to, interpreting, and responding to a child's explanation — each of these involves hearing and making sense of others' mathematical ideas in the moment, speaking on one's feet while seeking to connect with others. Although much of the work on mathematical knowledge needed by teachers is situated in relation to what teachers do, including using representations and interpreting students' thinking, as yet little of it has focused on the mathematical fluency needed for the work teachers do in classrooms, live, in communicating with students. As compellingly argued by Sfard (2008) and others (e.g., Resnick, Asterhan, & Clarke, 2015), it is this communicative work that is central to the practice of education. Failing to investigate and squarely address communicative mathematical demands of teaching may result in an impoverished theory of mathematical knowledge for teaching in ways that sorely limit its utility and impact.

Another major area of work centers on the need to address the persistent inequities in mathematics learning both produced and reproduced in school. Goffney and her colleagues have begun to identify a set of practices of equitable mathematics teaching (Goffney, 2010; Goffney & Gonzalez, 2015; Goffney, 2015), and several of the articles in this volume explore the measurement of mathematical knowledge for equitable teaching. The driving question is what do teachers need to appreciate and understand about mathematics in order to be able to create access for groups that have been historically marginalized? Part of this has to do with a flexible understanding of the mathematics that enables teachers to build bridges between mathematics and their

students. One aspect of this is to represent mathematics in ways that connect with their students' experience. Another is to be able to recognize mathematical capability and insight in their students' out-of-school practices. Each of these entails a flexibility of mathematical understanding, particularly of mathematical structure and practice. But it also involves the ability to recognize as mathematical a range of specific activities, reasoning processes, and ways of representing. Being able to do this can enable teachers to broaden both what it means to be "good at math" as well as what can be legitimated as "mathematics."

Equity is not a new focus in mathematics education (e.g., Schoenfeld, 2002), and there have been studies on the effect of gender and language on mathematics teachers' knowledge (Blömeke, Suhl, & Kaiser, 2011) as well as the distribution of teacher knowledge in different populations of teachers (Hill, 2007). In our review of the literature, we observed that most studies on equity were focused on aspirations and imperatives (i.e., arguments for teaching for equity). Few studies focused directly on specific practices of equitable mathematics teaching or knowledge for equitable mathematics teaching. We argue that increased focus in this area is crucial for three reasons. First is the underlying principle that extant inequity in mathematics teaching and learning is morally reprehensible in a civilized society (Perry, Moses, Cortez, Delpit, & Wynne, 2010). Second is our contention that, while certainly not in itself a solution, teacher content knowledge is both an indispensable and an untapped resource for disrupting the historical privileging of particular forms of mathematical competence and engagement. Third, as with nearly all achievement measures in early stages of development, current instruments are significantly biased because of the contextual features of where, as well as for and by whom, they are developed. The field needs good instrumentation, for research and for practice. Overly delaying the development of unbiased instruments may well undermine the political viability of well-meaning efforts to improve the mathematical education of teachers. Such development will require solid research in this difficult yet important arena.

In proposing these two areas of study, we acknowledge the conceptual and methodological challenges each presents. We suspect that research in these areas has been underdeveloped in large part because these foci involve subtle social dynamics less readily captured in print and in more conventional measures. These challenges simply add to our concern that concerted attention be given them. Our argument here is that these two areas of study are not merely our favored topics, but that they are essential to long-term success.

Articles that Develop Measures and Measure Development

The agenda sketched above is both a reflection of emerging work in the field and a proposal for future work. In many ways, the articles in this special issue, though specifically addressing measurement, resonate with themes above. For instance, the discussion about focused studies that contribute to a larger research program suggests some benefits of creating a common framework for describing mathematics teaching. In their article, Selling, Garcia and Ball (this issue) present a framework for unpacking the mathematical work of teaching that is promising in this respect. Whereas other frameworks often start with what teachers do, they focus first on the mathematical objects involved in the work of teaching and then follow up by describing actions that teachers

do on these objects. This idea builds upon and extends the notion of mathematical tasks of teaching that has been highlighted in previous publications on the practice-based theory of mathematical knowledge for teaching (e.g., Ball et al., 2008; Hoover, Mosvold, & Fauskanger, 2014), as well as in previous efforts to conceptualize the work of teaching (e.g., Ball & Forzani, 2009). A main aim with this framework is to inform and assist future development of items and instruments for measuring mathematical knowledge for teaching.

Phelps and Howell (this issue) discuss the role of teaching contexts in items developed to measure mathematical knowledge for teaching. Given that mathematical knowledge for teaching is understood as knowledge applied in the work of teaching, a teaching context that illustrates a certain component of this work is typically included in items. Phelps and Howell discuss different ways in which context can be critical to assessing mathematical knowledge for teaching. They argue that attention to the role of context might provide better understanding of the knowledge assessed in particular items and might also inform further development of a theory in which teaching context is used to define knowledge.

Whereas both of these first articles point to core issues regarding the conceptualization of mathematical knowledge for teaching — in the context of item and instrument development — the next two articles deal more directly with measurement. Kim (this issue) focuses on designing interview prompts for assessing mathematical knowledge for teaching. In particular, her discussion focuses on the task of providing written feedback to students. To model this task, she combines a decomposition of the task with aspects of the pedagogical context involved and sub-domains of mathematical knowledge for teaching.

Where Kim's study is more qualitative and conceptual in nature, Orrill and Cohen (this issue) draw on psychometric models in their work. Their study hinges on the issue of defining the construct measured, and they use a mixture Rasch model to analyze different subsets of items to support an argument that the domain definition has strong implications on the claims one tries to make about teachers' performance. In light of our observations about the lack of consensus about how to define the constructs that are being measured and discussed across studies, a focus on careful construct definition and implications is particularly relevant.

In the international literature on teaching and learning, a focus on equity is prevalent. In research on mathematical knowledge for teaching, the discussion of knowledge for teaching equitable mathematics also receives some attention — although issues of equity and diversity have not been emphasized in frameworks of mathematical knowledge for teaching. In this connection, Wilson's (this issue) and Turkan's studies of mathematical knowledge for teaching English language learners draw attention to this missing area of research. Both involve design and application of measures. Wilson proposes a new aspect of pedagogical content knowledge that is connected specifically to the work of teaching mathematics to English language learners. Turkan addresses practicing teachers' reasoning about teaching mathematics to ELLs. Based on analysis of data from cognitive interviews, she argues that there is a unique domain of knowledge

necessary for teaching ELLs — thus supporting Wilson's argument — and she calls for further investigations to identify and assess this knowledge.

Finally, this special issue includes two articles that investigate teachers' views. Koponen, Asikainen, Viholainen and Hirvonen investigate the views of teachers as well as teacher educators about the content of mathematics teacher education. Results from their survey indicate that teachers as well as teacher educators in the Finnish context emphasize the need for courses in content knowledge that is distinctive for teaching — not just more advanced. They argue that the mathematical content of teacher education needs to be tightly connected to the mathematics being taught, and even pedagogical courses need to include knowledge connected with mathematics, in particular focusing on knowledge of teaching and learning of mathematics. In the last article of the special issue, Kazima, Jakobsen and Kasoka investigate Malawian teachers' views about mathematical tasks of teaching and the potential usefulness of adapted measures of mathematical knowledge for teaching among Malawian pre-service mathematics teachers. The measures as well as the applied framework of mathematical knowledge for teaching were developed in the United States. Despite the significant cultural differences between Malawi and the United States, the authors argue that the framework as well as most of the items function well in the Malawian context.

Together, this collection of articles on the development and use of measures lies at a transition from the lessons of past studies of mathematical knowledge for teaching into vital arenas of research needed for systemic improvement on the mathematical education of teachers.

References

- Abbott, A. (1988). *The system of professions*. Chicago: University of Chicago Press.
- Adler, J., Ball, D., Krainer, K., Lin, F., & Novotna, J. (2008). Mirror images of an emerging field: Researching mathematics teacher education. In M. A. Niss (Ed.), *ICME-10 Proceedings: Proceedings of the 10th International Congress on Mathematical Education, 4-11 July 2004* (pp. 123–139). Roskilde: Roskilde Universitet.
- Akkerman, S. F., & Bakker, A. (2011). Boundary crossing and boundary objects. *Review of Educational Research*, 81(2), 132–169.
- American Educational Research Association. (2006). Standards for reporting on empirical social science research in AERA publications. *Educational Researcher*, 35(6), 33–40.
- Ball, D. L. (1999). Crossing boundaries to examine the mathematics entailed in elementary teaching. In T. Lam (Ed.), *Contemporary mathematics* (pp. 15–36). Providence, RI: American Mathematical Society.
- Ball, D. L., & Bass, H. (2003). Toward a practice-based theory of mathematical knowledge for teaching. In B. Davis & E. Simmt (Eds.), *Proceedings of the 2002 annual meeting of the Canadian Mathematics Education Study Group* (pp. 3–14). Edmonton, Alberta, Canada: Canadian Mathematics Education Study Group (Groupe Canadien d'étude en didactique des mathématiques).
- Ball, D. L., & Cohen, D. K. (1999). Developing practice, developing practitioners: toward a practice-based theory of professional education. In G. Sykes, & L. Darling-Hammond (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 3–32). San Francisco: Jossey Bass.
- Ball, D. L., & Forzani, F. M. (2009). The work of teaching and the challenge for teacher education. *Journal of Teacher Education*, 60(5), 497–511.
- Ball, D. L., Lubienski, S. T., & Mewborn, D. S. (2001). Research on teaching mathematics: the unsolved problem of teachers' mathematical knowledge. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed., pp. 433–456). New York, NY: Macmillan.
- Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, 59(5), 389–407.
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., ..., Tsai, Y. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, 47(1), 133–180.
- Begle, E. (1979). *Critical variables in mathematics education: Findings from a survey of empirical research*. Mathematical Association of America/National Council of Teachers of Mathematics, Washington, DC.
- Bell, C. A., Wilson, S. M., Higgins, T., & McCoach, D. B. (2010). Measuring the effects of professional development on teacher knowledge: The case of developing

- mathematical ideas. *Journal for Research in Mathematics Education*, 41(5), 479–512.
- Blömeke, S., & Delaney, S. (2012). Assessment of teacher knowledge across countries: a review of the state of research. *ZDM*, 44(3), 223–247.
- Blömeke, S., Suhl, U., & Kaiser, G. (2011). Teacher education effectiveness: Quality and equity of future primary teachers' mathematics and mathematics pedagogical content knowledge. *Journal of Teacher Education*, 62(2), 154–171.
- Boerst, T. A., Sleep, L., Ball, D. L., & Bass, H. (2011). Preparing teachers to lead mathematics discussions. *Teachers College Record*, 113(12), 2844–2877.
- Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational researcher*, 33(8), 3–15.
- Charalambous, C. (2010). Mathematical Knowledge for Teaching and Task Unfolding: An Exploratory Study. *The Elementary School Journal*, 110(3), 247–278.
- Chick, H. (2009). Choice and Use of Examples as a Window on Mathematical Knowledge for Teaching. *For the Learning of Mathematics*, 29(3), 26–30.
- Cohen, D., Raudenbush, S., & Ball, D. L. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis*, 25(2), 1–24.
- Copur-Gencturk, Y., & Lubienski, S. T. (2013). Measuring mathematical knowledge for teaching: a longitudinal study using two measures. *Journal of Mathematics Teacher Education*, 16(3), 211–236.
- Depaepe, F., Verschaffel, L., & Kelchtermans, G. (2013). Pedagogical content knowledge: a systematic review of the way in which the concept has pervaded mathematics educational research. *Teaching and Teacher Education*, 34, 12–25.
- Elliott, R., Kazemi, E., Lesseig, K., Mumme, J., & Kelley-Petersen, M. (2009). Conceptualizing the Work of Leading Mathematical Tasks in Professional Development. *Journal of Teacher Education*, 60(4), 364–379.
- Ferguson, P., & Womack, S. T. (1993). The impact of subject matter and education coursework on teaching performance. *Journal of Teacher Education*, 44, 155–163.
- Gherardi, S. (2012). *How to conduct a practice-based study: Problems and methods*. Northampton, MA: Edward Elgar Publishing.
- Goffney, I. M. (2010). *Identifying, defining, and measuring equitable mathematics instruction*. (Unpublished doctoral dissertation). University of Michigan, Ann Arbor, MI.
- Goffney, I. M., & Gonzalez, M. (2015). Mathematics methods as a site for developing ambitious and equitable teaching practices for prospective teachers. Manuscript submitted for publication.
- Goffney, I. M. (2015). Developing a theory of mathematical knowledge for equitable teaching. Manuscript in preparation.

- Graeber, A., & Tirosh, D. (2008). Pedagogical content knowledge: Useful concept or elusive notion. In P. Sullivan & T. Woods (Eds.), *Knowledge and beliefs in mathematics teaching and teaching development* (pp. 117–132). Rotterdam: Sense Publishers.
- Guyton, E., & Farokhi, E. (1987). Relationships among academic performance, basic skills, subject matter knowledge, and teaching skills of teacher education graduates. *Journal of Teacher Education*, 38, 37–42.
- Harbison, R.W., & Hanushek, E. A. (1992). *Educational performance for the poor: Lessons from rural northeast Brazil*. Oxford, UK: Oxford University Press.
- Herbst, P., & Kosko, K. (2012). Mathematical knowledge for teaching high school geometry. In L. R. Van Zoest, J.-J. Lo, & J. L. Kratky (Eds.), *Proceedings of the 34th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 438–444). Kalamazoo, MI: Western Michigan.
- Hill, H. (2007). Mathematical knowledge of middle school teachers: implications for the no child left behind policy initiative. *Educational Evaluation and Policy Analysis*, 29(2), 95–114.
- Hill, H. C. (2011). The nature and effects of middle school mathematics teacher learning experiences. *Teachers College Record*, 113(1), 205–234.
- Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction*, 26(4), 430–511.
- Hill, H. C., Kapitula, L., & Umland, K. (2011). A validity argument approach to evaluating teacher value-added scores. *American Educational Research Journal*, 48(3), 794–831.
- Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Education Research Journal*, 42(2), 371–406.
- Hill, H. C., Schilling, S. G., & Ball, D. L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *The Elementary School Journal*, 105(1), 11–30.
- Hill, H. C., Umland, K., Litke, E., & Kapitula, L. (2012). Teacher quality and quality teaching: Examining the relationship of a teacher assessment to practice. *American Journal of Education*, 118(4), 489–519.
- Hoover, M., Mosvold, R., & Fauskanger, J. (2014). Common tasks of teaching as a resource for measuring professional content knowledge internationally. *Nordic Studies in Mathematics Education*, 19(3–4), 7–20.

- Izsak, A., Jacobson, E., de Araujo, Z., & Orrill, C. H. (2012). Measuring mathematical knowledge for teaching fractions with drawn quantities. *Journal for Research in Mathematics Education*, 43(4), 391–427.
- Jacobson, E., Remillard, J., Hoover, M., & Aaron, W. (in press). Chapter 8: The interaction between measure design and construct development: Building validity arguments. *Journal for Research in Mathematics Education. Monograph*.
- Kaarstein, H. (2014). A comparison of three frameworks for measuring knowledge for teaching mathematics. *Nordic Studies in Mathematics Education*, 19(1), 23–52.
- Kersting, N. B., Givvin, K. B., Sotelo, F. L., & Stigler, J. W. (2010). Teachers' analyses of classroom video predict student learning of mathematics: further explorations of a novel measure of teacher knowledge. *Journal of Teacher Education*, 61, 172–181.
- Kersting, N. B., Givvin, K. B., Thompson, B. J., Santagata, R., & Stigler, J. W. (2012). Measuring usable knowledge teachers' analyses of mathematics classroom videos predict teaching quality and student learning. *American Educational Research Journal*, 49(3), 568–589.
- Kim, Y., Mosvold, R., & Hoover, M. (2015). A review of research on the development of teachers' mathematical knowledge for teaching. Manuscript in preparation.
- Kleickmann, T., Richter, D., Kunter, M., Elsner, J., Besser, M., Krauss, S., & Baumert, J. (2013). Teachers' content knowledge and pedagogical content knowledge: The role of structural differences in teacher education. *Journal of Teacher Education*, 64(1), 90–106.
- Koellner, K., Jacobs, J., Borko, H., Schneider, C., Pittman, M. E., Eiteljorg, E., Bunning, K., & Frykholm, J. (2007). The Problem-Solving Cycle: A model to support the development of teachers' professional knowledge. *Mathematical Thinking and Learning*, 9(3), 271–300.
- Kunter, M., Klusmann, U., Baumert, J., Voss, T., & Hacfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology*, 105(3), 805–820.
- Lortie, D. C. (1975). *School teacher: A sociological inquiry*. Chicago, IL: University of Chicago Press.
- Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Mahwah, NJ: Lawrence Erlbaum Associates.
- McCray, J. S., & Chen, J.-Q. (2012). Pedagogical content knowledge for preschool mathematics: Construct validity of a new teacher interview. *Journal of Research in Childhood Education*, 26, 291–307.
- McCrory, R., Floden, R., Ferrini-Mundy, J., Reckase, M. D., & Senk, S. L. (2012). Knowledge of algebra for teaching: A framework of knowledge and practices. *Journal for Research in Mathematics Education*, 43(5), 584–615.

- Monk, D. H. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of Education Review*, 13, 125–145.
- Morris, A. K., Hiebert, J., & Spitzer, S. (2009). Mathematical knowledge for teaching in planning and evaluating instruction: What can preservice teachers learn? *Journal for Research in Mathematics Education*, 40(5), 491–529.
- Mullens, J. E., Murnane, R. J., & Willett, J. B. (1996). The contribution of training and subject matter knowledge to teaching effectiveness: A multilevel analysis of longitudinal evidence from Belize. *Comparative Education Review*, 40, 139–157.
- Perry, T., Moses, R. P., Cortes Jr, E., Delpit, L., & Wynne, J. T. (2010). *Quality education as a constitutional right: Creating a grassroots movement to transform public schools*. Boston, MA: Beacon Press.
- Resnick, L. B., Asterhan, C., & Clarke, S. N. (Eds.). (2015). *Socializing intelligence through academic talk and dialogue*. Washington, D.C.: American Educational Research Association.
- Rockoff, J. E., Jacob, B. A., Kane, T. J., & Staiger, D. O. (2011). Can you recognize an effective teacher when you recruit one? *Education Finance and Policy*, 6(1), 43–74.
- Saderholm, J., Ronau, R., Brown, E. T., & Collins, G. (2010). Validation of the Diagnostic Teacher Assessment of Mathematics and Science (DTAMS) instrument. *School Science and Mathematics*, 110(4), 180–192.
- Schilling, S. G., & Hill, H. C. (2007). Assessing measures of mathematical knowledge for teaching: a validity argument approach. *Measurement*, 5(2–3), 70–80.
- Schoenfeld, A. H. (2002). Making mathematics work for all children: Issues of standards, testing, and equity. *Educational researcher*, 31(1), 13–25.
- Schwab, J. J. (1978). Education and the structure of the disciplines. In I. Westbury & N. Wilkof (Eds.), *Science, curriculum, and liberal education: Selected essays* (pp. 229–272). Chicago, IL: University of Chicago Press. (Original work published 1961.)
- Senk, S. L., Tatto, M. T., Reckase, M., Rowley, G., Peck, R., & Bankov, K. (2012). Knowledge of future primary teachers for teaching mathematics: An international comparative study. *ZDM*, 44(3), 307–324.
- Sfard, A. (2008). *Thinking as communicating: Human development, development of discourses, and mathematizing*. Cambridge: Cambridge University Press
- Shulman, L. S. (1986). Those who understand: knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14.
- Silver, E. A., Clark, L. M., Ghouseini, H. N., Charalambous, C. Y., & Sealy, J. T. (2007). Where is the mathematics? Examining teachers' mathematical learning opportunities in practice-based professional learning tasks. *Journal of Mathematics Teacher Education*, 10(4–6), 261–277.

- Speer, N. M., & Wagner, J. F. (2009). Knowledge needed by a teacher to provide analytic scaffolding during undergraduate mathematics classroom discussions. *Journal for Research in Mathematics Education*, 40(5), 530–562.
- Star, S. L., & Griesemer, J. R. (1989). Institutional ecology, “translations” and boundary objects: Amateurs and professionals in Berkeley’s Museum of Vertebrate Zoology, 1907–39. *Social Studies of Science*, 19, 387–420.
- Stein, M. K., Engle, R. A., Smith, M. S., & Hughes, E. K. (2008). Orchestrating productive mathematical discussions: Five practices for helping teachers move beyond show and tell. *Mathematical Thinking and Learning*, 10(4), 313–340.
- Tatto, M. T., Schwille, J., Senk, S., Ingvarson, L., Peck, R., & Rowley, G. (2008). *Teacher Education and Development Study in Mathematics (TEDS-M): Conceptual framework*. East Lansing, MI: Teacher Education and Development International Study Center, College of Education, Michigan State University.
- Thames, M. H. (2009). *Coordinating mathematical and pedagogical perspectives in practice-based and discipline-grounded approaches to studying mathematical knowledge for teaching (K-8)*. (Unpublished doctoral dissertation). University of Michigan, Ann Arbor, MI.
- Thompson, P. W. (2015). Researching mathematical meanings for teaching. In L. D. English & D. Kirshner (Eds.), *Third handbook of international research in mathematics education* (pp. 435–461). New York: Taylor & Francis.
- Williams, S. R., & Baxter, J. A. (1996). Dilemmas of discourse-oriented teaching in one middle school mathematics classroom. *The Elementary School Journal*, 97, 21–38.
- Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation and autonomy in mathematics. *Journal for Research in Mathematics Education*, 27, 458–476.
- Yongs, P., & Qian, H. (2013). The influence of university courses and field experiences on Chinese elementary candidates’ mathematical knowledge for teaching. *Journal of Teacher Education*, 64(3), 244–261.
- Zodik, I., & Zaslavsky, O. (2008). Characteristics of teachers’ choice of examples in and for the mathematics classroom. *Educational Studies in Mathematics*, 69(2), 165–182.

What Does it Take to Develop Assessments of Mathematical Knowledge for Teaching?:

Unpacking the Mathematical Work of Teaching

Sarah Kate Selling
University of Michigan

Nicole Garcia
University of Michigan

Deborah Loewenberg Ball
University of Michigan

Abstract: In the context of the increased mathematical demands of the Common Core State Standards and data showing that many elementary school teachers lack strong mathematical knowledge for teaching, there is an urgent need to grow teachers' MKT. With this goal in mind, it is crucial to have research and assessment tools that are able to measure and track aspects of teachers' MKT at scale. Building on the concept of “mathematical tasks of teaching” (Ball et al., 2008), we report on a new framework that unpacks the mathematical work of teaching that could serve as a scaffold for item writers who are developing assessments of MKT. We argue that this framework supports a focus on the *mathematical* work of teaching that moves beyond common content knowledge but without moving into a space of pedagogical choice. We also illustrate how the framework was constructed to highlight connections within and across the mathematical content of elementary school. The mathematical work of teaching framework has implications for assessment development at scale, and could be useful as an organizing tool in mathematics teacher education efforts to grow teachers' MKT.

Keywords: mathematical knowledge for teaching, teacher knowledge, assessment development

Introduction

Broad consensus exists about the importance of teachers' mathematical knowledge (Adler & Venkat, 2014; Ball, Lubienski, & Mewborn, 2001; Baumert et al. 2010; Döhrmann, Kaiser, & Blömeke, 2014). Studies have linked mathematical knowledge for teaching to the quality of teachers' mathematics instruction (Eisenhart, Borko, Underhill, Brown, Jones, & Agard, 1993; Hill et al., 2008). Mathematical knowledge for teaching has also been linked to

student achievement gains in the elementary grades (Hill, Rowan, and Ball, 2005). However, many U.S. teachers lack the deep, nuanced, and specialized mathematical knowledge needed for responsible teaching. This finding is persistent over time, grade levels, and both national and international contexts (e.g., Hill & Ball, 2004; Ma, 1999; Tatto et al., 2008). Simultaneously, the Common Core State Standards for Mathematics (National Governors Association Center for Best Practices, 2010), which have been adopted by 47 states and territories, have set out rigorous standards for K-12 mathematics learning that consequently increase the mathematical demands of teaching. To ensure that teachers are well-positioned to help students meet these more challenging learning goals, it is now — more than ever — critically important to focus on developing their mathematical knowledge for teaching.

To investigate and grow teachers' mathematical knowledge for teaching (MKT) it is crucial to be able to measure and track the development and uses of MKT. Most work to develop measures of MKT has typically been done by groups of experts in relevant fields, such as mathematicians, mathematics educators, and teachers who have worked together to draft and revise assessment items (Hill, Schilling, & Ball, 2004). The early work in this area was focused on developing and refining the construct of mathematical knowledge for teaching while simultaneously and iteratively developing measures of the construct. The process of item development was therefore often time consuming and challenging. Because of the promising results of these earlier efforts, there is now a broad need for assessments of MKT. Building tests at scale means, however, that people who are not deeply immersed in research on MKT will have to be able to write valid MKT items. This will require detailed supports to help test developers understand the nuances of the construct of MKT and ways to assess it. In this paper, we present a framework that identifies the different ways that teachers make use of mathematical knowledge as they go about the work of teaching and provides support to assessment developers. We begin by articulating and specifying what we mean by mathematical knowledge for teaching and its relationship with the mathematical work of teaching that arises in everyday practice.

Theoretical Framing

Conceptualizing Mathematical Knowledge for Teaching

Building supports for assessment development of mathematical knowledge for teaching (MKT) rests on a clear conceptualization of what we mean by MKT, how MKT is drawn upon in practice, and the specific areas of the work of teaching that we seek to assess. Scholars of mathematical knowledge have examined such knowledge in action as it is used in the practice of teaching (Ball & Bass, 2002; Rowland, Huckstep, & Thwaites, 2005).

Our work builds on a particular practice-based perspective on mathematical knowledge for teaching that begins with the premise that, to understand the specific knowledge of mathematics needed in teaching, one must first examine the mathematical work that arises in the context of teachers' instruction in classrooms, a form of job analysis (Ball & Bass, 2002). Through detailed analysis of instruction in a 3rd grade classroom over an entire year, Ball and her colleagues identified mathematical problems that teachers regularly encounter and must solve while teaching, such as “interpreting and evaluating students' non-standard mathematical ideas” (Ball & Bass, 2002, p. 9). These analyses reveal that teaching entails significant mathematical work on the part of the teacher. To highlight the complexity and variety of ways that teachers engage in mathematical work, Ball and her colleagues (2008) present a list of 16 “mathematical tasks of teaching” that may occur within every day teaching practice that involve mathematical

work on the part of the teacher. This list includes tasks such as “responding to students’ ‘why’ questions”, “finding an example to make a specific mathematical point”, “evaluating the plausibility of students’ claims (often quickly)”, “choosing and developing useable definitions”, or “recognizing what is involved in using a particular representation” (p. 10). These mathematical tasks of teaching provide the contexts in which teachers must draw on mathematical knowledge for teaching, and therefore offer a window into the mathematical knowledge entailed by teaching.

Based on their analyses of these ubiquitous tasks of teaching mathematics, Ball and her colleagues (2008) identified a provisional map of domains of mathematical understanding and skill. They argued that teaching requires both “pure” subject matter knowledge and pedagogical content knowledge (Shulman, 1986, 1987; Wilson, Shulman, and Richert, 1987). Pedagogical content knowledge comprises blends of mathematical knowledge together with other kinds of knowledge, such as knowledge of students’ thinking in a particular content domain, or knowledge of likely effective approaches to or materials for teaching specific content ideas. For example, in teaching integers, teachers need to appreciate that notions of “debt,” “assets” and “net worth” are unfamiliar to elementary age learners and that therefore financial contexts are not likely to be useful as a representation of integer arithmetic. Knowing ways to use number line models as a context for integer arithmetic is another example of pedagogical content knowledge — knowledge of teaching approaches and models combined with a particular topic. But knowing integers for teaching also involves content knowledge. “Common” content knowledge is the term Ball and her colleagues use to describe the knowledge that 0 is neither negative or positive or that $(-3) - (-7) = 4$. By this they denote knowledge that is also relevant to people who do not teach — that is, known *in common* with others. They argue that teaching also requires “specialized” content knowledge — for example, being able to explain the meaning of subtraction of a negative number and connect it to moves on the number line in ways that make conceptual sense, or being able to represent the difference – even though they might produce the same result — between subtracting -4 from 10 and adding 4 to 10. Horizon knowledge is the perspective needed to understand connections among topics or to see where ideas are headed, or to notice when students are onto a sophisticated mathematical point (Ball & Bass, 2009). In our assessment development work, we focus on specialized content knowledge, as a form of subject matter knowledge that is particularly needed in the work of teaching.

Figure 1. Domains of Content Knowledge for Teaching (Ball, Thames, & Phelps, 2008)

Research on specialized content knowledge has acknowledged that the line between specialized and common content knowledge might not be well-defined, and that particular

mathematical tasks of teaching may elicit different types of knowledge by teachers or others asked to engage in these tasks (Delaney et al., 2005; Hill, Dean, & Goffney, 2007). In our work, we are less concerned with classifying assessment items as eliciting only specialized or common content knowledge; instead, we have chosen to focus on the mathematical work of teaching demanded by teaching practice and the knowledge that teachers would need to do that work, acknowledging that some mathematical work of teaching may elicit different domains of subject matter knowledge or even knowledge from multiple domains.

Building Assessments of Content Knowledge for Teaching: Challenges and Supports

Existing assessments of teacher knowledge at scale, often licensure tests, tend to focus on common content knowledge (i.e., the mathematics content that teachers teach) or horizon knowledge (i.e., perspective on how what the students are working on now connects with other mathematics). Few assessments have attempted to assess specialized content knowledge at scale. The Learning Mathematics for Teaching project (Hill & Ball, 2004; Hill, Schilling & Ball, 2004) has developed elementary and middle school level measures for research purposes that have been widely adopted and implemented. However, these measures are not intended as assessments of individual teachers. This leaves unaddressed how to develop assessments of SCK at scale and how to support item writing by test designers. Our investigation of this question has been situated in a project in which we collaborated with others to build items to measure teachers' specialized content knowledge at scale. Our goal was to develop tools that could be used to guide the development of assessments of SCK with item writers who have different expertise than the groups who have in the past worked to develop items like those in the Learning Mathematics for Teaching project.

To understand what tools and supports might be needed to accomplish this, we considered what might be challenging for assessment developers when constructing measures of specialized content knowledge. First, we hypothesized that item writers might have difficulty developing measures of more than just common content knowledge, as this is the typical focus for assessments of teacher knowledge. In particular, we anticipated that there would be challenges in understanding the differences between CCK and SCK. A second related challenge concerns the possibility that in attempting to shift from writing items focused on common content knowledge, item writers might end up going too far and focusing items on pedagogical tasks of teaching that involve more than mathematical work, such as making instructional decisions about the best ways to teach a topic. In other words, we were concerned that writers might develop items focused on pedagogical content knowledge or even pedagogical choices, both which were beyond the scope of a subject matter knowledge for teaching assessment. A third challenge might arise if item writers are not familiar with the work of teaching that draws on teachers' specialized mathematical knowledge, such as the tasks of teaching set out by Ball and colleagues (2008). Finally, we hypothesized that it might also be difficult for item writers to understand how specialized content knowledge might be used across the K-6 curriculum, and how those uses might vary. Based on these four hypothesized areas of difficulty, we developed a framework that identifies the mathematical work of teaching and is strategically designed to address each of these challenges. We highlight below how the framework supports a focus on the *mathematical* work of teaching that moves beyond common content knowledge but without moving into a space of pedagogical choice. We also illustrate how the framework was constructed to highlight connections within and across the mathematical content of elementary school.

Unpacking the Mathematical Work of Teaching Framework

The mathematical work of teaching framework expands on the mathematical tasks of teaching (Ball et al., 2008) to produce a tool that can support development of assessments of mathematical knowledge for teaching at the elementary level. The framework addresses three main goals. First, the framework supports a focus on the *mathematical* work of teaching, the mathematics that a teacher engages with while teaching content to students, as opposed to the pedagogical task of making choices about instructional strategies. Second, the framework highlights connections between the mathematical work of teaching and the mathematics content at the elementary level. Finally, the framework is usable by item writers to construct written measures of MKT, specifically measures of subject matter knowledge with a focus on specialized content knowledge. In the following sections, we unpack the mathematical work of teaching framework with respect these three goals, referencing an excerpt from the framework shown below in Table 1.

Table 1: *Mathematical work of teaching framework organized by (1) mathematical objects, (2) actions with and on those objects in teaching, and (3) specific examples.*

	MWT: Actions with and on objects	Examples
Explanations (includes justifications & reasoning)	Comparing explanations to determine which is more/most valid, generalizable, or complete explanation	Given two explanations, choose which is more complete Given multiple student explanations, determine which is most valid Given several explanations, choose the best explanation. Given conflicting explanations, determine which is valid and why. Select an explanation that best captures an underlying idea.
	Critiquing explanations to improve them with respect to completeness, validity, or generalizability.	Given an incomplete but valid explanation, determine what, if anything, is missing or needs to be added to be more complete.
	Critiquing explanations with respect to validity, generalizability, or explanatory power.	Given an explanation, determine if it is mathematically valid. Given several explanations, determine which ones are valid. Given a text, determine what may be misleading about an explanation.
	Writing mathematically valid explanations for a process, conjecture, relationship, etc.	Write a mathematically valid explanation for a process or concept. Write a mathematically valid explanation for a conjecture. Given student strategies, determine properties that could be used to justify the strategy's validity.
Mathematical structure	Determining, analyzing, or posing problems with the same (or different) mathematical structure	Given a set of problems, determine which have the same structure. Given a set of problems, choose the description of the structure type. Given a set of problems, determine which does NOT have the same structure. Write a problem that has the same structure as given problems. Given a description of a structure, determine which problems fit that structure.
	Analyzing structure in student work by determining which strategies or ideas are most closely connected with respect to mathematical structure	Given a set of student strategies, determine which have similar mathematical structure. Given a set of student strategies most of which use the same core idea but slightly differently, determine which one does not fit. Given a set of strategies and a structure, determine which strategies fit the structure.
	Matching word problems and structure	Given a structure, choose a word problem with that structure. Given a word problem, choose another problem with the same structure.

Representations	Connecting or matching representations	Match a representation to a given interpretation of an operation. Determine how different representations are connected. Given two claims about a representation, determine which is correct and why.
	Analyzing representations by identifying correct or misleading representations in a text, talk, or written work.	Given a written representation (e.g., number line, table, diagram), determine what may be misleading. Given a set of representations, choose which does or does not show a particular idea (table?)
	Selecting, creating, or evaluating representations for a mathematical purpose	Create a representation for a given number or operation. Select a representation that highlights a particular mathematical idea.
	Talking a representation (i.e., using words to talk through the meaning of a representation and connecting it to the key ideas)	Given a suggested way to talk about a representation in a text, evaluate whether the talk clearly connects the representation and the ideas. Given a colleague's request for feedback, determine how their talking about a representation could be improved to highlight mathematical meaning

Organization around “Mathematical Objects”

Our first task in developing the mathematical work of teaching framework was to organize the mathematical work that arises in the context of teaching in a way that would maintain a focus on the mathematics. The mathematical tasks of teaching, as set out by Ball and colleagues (2008), include a list of 16 illustrative tasks that arise in everyday practice and that entail mathematical work for the teacher. This list includes tasks such as “responding to students’ ‘why’ questions”, “finding an example to make a specific mathematical point” or “recognizing what is involved in using a particular representation” (p. 10). These provide useful examples of the mathematical work of teaching; however, they were not intended to be an exhaustive list. Therefore, we built on and expanded this list. As the list of teachers’ mathematical work grew, we needed to create an organizational structure that would make the list more orderly, systematic, and useful for item writers.

The mathematical work of teaching framework is organized around a set of what we call “mathematical objects” that teachers encounter and with which they work while teaching. Examples include explanations, representations, mathematical errors, and definitions. We called these “mathematical objects” because they are the mathematical instructional objects that teachers encounter and with which they interact while teaching. For example, teachers regularly give, use, and encounter mathematical explanations; in this case, the mathematical explanation is the “object”. Teachers give mathematical explanations themselves, but they also make sense of student explanations, compare different explanations in textbooks, determine if a student’s explanation is valid, or critique written explanations for the purpose of improving them. We recognize that we define “mathematical objects” here in a way that is different from the way “objects” is typically used in mathematics to refer to objects such as numbers, functions, and polygons. Table 2 provides the set of mathematical objects around which we built the framework. Although this list is by no means exhaustive, we hoped to describe the diverse sets of mathematical objects that teachers typically interact with in teaching.

Table 2. *Organizing mathematical objects for the mathematical work of teaching.*

Explanations (including justification and reasoning)	Errors and incorrect thinking
Conjectures	Representations
Mathematical structure	Manipulatives
Examples, non-examples, and counter-examples	Language and definitions
Mathematical problems	Mathematical goals and topics
Strategies	

Establishing the framework around a set of mathematical objects focuses attention on the *mathematical* work of teaching. By basing the framework in these mathematical objects, the mathematics in teachers' work is foregrounded. Another way to organize such a framework would be to organize the mathematical work of teaching into pedagogical tasks or domains (e.g., the mathematical work that arises when leading a discussion). However, there are two main limitations in organizing the framework in this way. First, many of the mathematical tasks of teaching arise in the context of enacting multiple different instructional practices, but require the same mathematical work on the part of the teacher, regardless of context. For example, interpreting student mathematical errors, a key mathematical task of teaching, could arise in the context of a class discussion, but it might also arise while teachers are interpreting written work on assessments or during teachers' interactions with individuals or small groups of students. Representing the mathematical work of teaching in each of these instructional practices would result in a lengthy list of work with much repetition. A second reason to avoid organizing the framework around pedagogical domains or tasks is to keep the focus on the *mathematical* work of teaching to help item writers avoid developing items that simply assessed teachers' instructional choices. Organization around instructional practices emphasizes the teaching practice rather than the mathematics necessary to engage in that practice. Consider the instructional practice of giving oral or written feedback to students. This practice requires teachers to engage in mathematical work such as determining how a student's explanation could be improved to be more complete. Organizing the framework around mathematical objects, as opposed to instructional practices or other pedagogically focused categories, supports a focus on the mathematics in the work of teaching.

Organization around Mathematical Work of Teaching with Respect to these Objects

The framework is organized around a diverse set of mathematical objects to illuminate the varied mathematical terrain of teachers' work, ranging from the mathematics that arises in interacting with explanations and strategies, to the mathematics involved in using language and definitions carefully, to the mathematical work of choosing or constructing mathematical examples. Explicitly naming and building the framework around this set of objects highlights the diversity of mathematical work of teaching.

Each domain of mathematical objects is further defined by a set of mathematical work of teaching, or actions on that particular object. For example, the mathematical object of "representations" includes five categories of work, i.e., "connecting or matching representations", "analyzing representations", "choosing or creating representations", and

“talking a representation”. The categories of work for each object answer the question, “what mathematical work do teachers do with these objects while teaching”? The grain-size of these categories manages the tension between (1) defining a useful set of categories that adequately captures the nuance and variability of the ways in which teachers interact with these objects, and (2) constructing a list that is manageable to use rather than a long list of specific verbs and scenarios. These categories comprise a larger domain of mathematical work that could be defined by introducing different mathematical criteria. For example, explanations, as objects, could be critiqued by teachers with respect to the validity, generalizability, or completeness of the explanations. Rather than define a separate category for each type of critique, the framework groups them as one category of mathematical work of teaching around the mathematical object of explanations. Each category might also refer to a range of contexts in which teachers might engage in this particular mathematical work. For example, as shown in Table 1 the category of analyzing representations includes reference to a set of contexts in which teachers might encounter representations, such as written work, talk, or texts. This organization of the framework keeps the set of categories concise while also mapping the dimensions of variation in teachers’ mathematical work. This could support item writers in sampling across the varied terrain of teacher’s mathematical work. Although the framework is detailed, it is not fully intended to represent all of the work of teaching, but to focus on common ways that teachers interact with particular objects.

It is important to note that the high-level categories built around mathematical objects are not necessarily disjoint. We utilize the focus of the mathematical work to determine where particular tasks of teaching fit in the framework. For example, consider the mathematical work of teaching that involves analyzing student strategies in written work for evidence of use of a particular mathematical structure (e.g., examining whether strategies reveal evidence of a comparison or take-away interpretation of subtraction). This work involves both student strategies and mathematical structure. However, the framework classifies the work of teaching by the main mathematical focus or goal, so looking for structure in student strategies would be classified as belonging to mathematical structure because looking for mathematical structure is the primary mathematical work, while the student strategy was the context in which it arose.

The third level of the framework further illustrates each category of work with examples that could serve as “shells” for items or “item starters.” These examples do not include all of the necessarily details that might exist in a finished item, but could serve as a starter for beginning item developer to write items in this category. The examples are specific enough to provide help in beginning to write an item in that category. Consider the category of critiquing explanations. One example in that category is “Given several explanations, determine which ones are valid.” This includes information about key elements that would need to be specified in the item (i.e. several explanations) and provides the desired action on the part of the test taker (i.e. determining validity). Each category includes several examples to help item writers attend to the different contexts and ways the work might play out, but the framework makes clear that the set of examples is not exhaustive and there are other ways to construct items and scenarios in each category.

Interactions with Mathematical Content

The framework also provides support in understanding the mapping between particular K-6 mathematical content and the mathematical work of teaching. For example, the framework helps answer questions such as “in which mathematical content areas do teachers most likely

interact with mathematical definitions”? Although the mathematical work of teaching can be mapped to all K-6 mathematics content, the framework focuses the supports on the most critical, or high-leverage K-6 topics (Ball & Foran, 2011). The mathematical work of teaching interacts with this content in a number of ways. First, there are some categories of work with respect to particular mathematical objects that are likely to emerge in instruction across all content areas. For example, interpreting students’ mathematical errors is part of the mathematical work of teaching all mathematical topics (e.g., number and operations, measurement, fractions). In contrast, there are other parts of the mathematical work of teaching that are more likely to arise when teaching particular mathematical content. Consider “critiquing strategies”, a category of teacher’s work. While certainly possible that teachers might engage in this work across all content areas, there are some content areas in which teacher might need to do this work more frequently and with a set of common strategies. For example, when teaching multi-digit subtraction, teachers are likely to have to analyze and critique students’ non-standard strategies. Similarly, making sense of student strategies is also likely to be part of the work when students are learning to compare fractions when there are many common strategies for doing so (e.g., common numerators, benchmarking). In contrast, this type of work is less likely to emerge when teaching aspects of geometry. The framework serves as a scaffold for item writers to think about in which mathematical content particular work is most likely to happen.

Annotations are included in the fourth column of the framework to foreground these connections and to highlight other considerations for writing items. These annotations address multiple areas of concern for item writers, often suggesting mathematical topics that are a good fit (or less good fit) with that category of work. In the case of the critiquing strategies example described above, the framework includes a note that “ordering numbers, operations with numbers” are fruitful areas for writing items in this category. Other times, the framework indicates that all content areas are a good fit. This column also includes annotations about the challenges of writing items in certain categories and with particular content. For example, in the topic of comparing fractions, there are a number of strategies that will result in the correct answer in some but not all cases, which makes this a productive terrain for writing items that assess candidates’ ability to critique the validity and generalizability of strategies. In contrast, with whole number operations, it is much more difficult to find examples of strategies that either only work for a subset of whole numbers or strategies that result in a correct answer but are not valid. Developing items in this space is therefore quite challenging and requires very careful and strategic selection of numbers. To help item writers understand this interaction between the MWT category and this content, the fourth column includes a note to indicate this difficulty.

This fourth column provides additional varied supports for understanding the MWT framework, including identifying potential item types and interactions with content. For example, some categories of the framework are areas in which others have developed items that serve as model items noted in the annotations, whereas others are novel in the sense that very few (or even no) examples of items in that space exist. The framework includes these annotations to describe the range of work teachers do and to inspire the development of new types of items that assess this range. However, items in this space are likely to be more difficult to write without examples from which to build. Therefore, notes in the fourth column alert item writers to the fact that the category was new and potentially challenging to write to. An example of this is the category of “Critiquing the use of a representation”, meant to capture the work that teachers might need to do when making sense of the use of representations in particular ways by students, other teachers, or curricular materials. The final type of annotation in the framework consists of

notes to support item writers in maintaining the focus on the mathematical work and refraining from building items that focus on pedagogical choices. Early observations of item writing indicate that certain categories of work (e.g., manipulatives, errors) are more likely to lead item writers into pedagogical terrain, such as presenting a student error and asking the candidate to decide the next step that would best help that student or choosing the best manipulative to help a child see his or her mistake. Annotations are included to alert item writers when a particular category of work provides challenge for retaining the focus on the mathematics, along with common mistakes made in writing items too focused on pedagogy.

Utilizing the Framework for Item Development

To further specify the use of the mathematical work of teaching framework, we will examine how the table can be used to develop items, with a focus on explanations and mathematical structure. We begin by selecting the mathematical object of explanations and one key piece of the mathematical work of teaching with that object, “critiquing explanations with respect to validity, generalizability, or explanatory power.” For this example, we will focus on the criteria of generalizability. To develop an item focused on the mathematical work a teacher does when critiquing explanations for generalizability, we must consider what mathematics content makes available a variety of explanations that may or may not be generalizable. One area of mathematics that is ripe with both explanations and methods that may or may not be generalizable is numbers and operations. For this example, we focus on operations with decimals, specifically decimal multiplication. The item shown in Table 3 requires a teacher to determine for each given explanation, whether or not the explanation represents generalizable methods for multiplying any two decimals. This is mathematical work that teachers do on a regular basis in a variety of contexts.

The second sample item is focused on the mathematical work of teaching involved in creating problems with a particular mathematical structure. In this case, the selected object for the item is “mathematical structure” and is combined with the work of “determining, analyzing, or posing problems with the same (or different) mathematical structure.” Again, we must consider the mathematics content that teachers are most likely to encounter the need to determine or highlighting the mathematical structure of the work. Division is one area of elementary mathematics where particular interpretations of the operation require teachers to attend to mathematical structure. The problem shown in Table 3 requires one to apply a measurement (or quotitive) interpretation of division to develop a word problem. This involves careful attention to the structure of measurement division problems with attention to the meaning of each of the parts of the problems and then transferring this meaning to a particular context.

Table 3. *Examples of items written at the intersection of the mathematical work of teaching and content.*

Explanations Sample Item	
Mr. Reinke is working with his students on decimal multiplication. He asked them to solve the problem 1.2×0.3 and explain their process.	
Which of the following student explanations for multiplying 1.2×0.3 represent methods that are generalizable for multiplying any two decimal numbers? Select <u>all</u> that apply.	
(A) "I just multiplied 3×12 and got 36. I counted the total numbers behind the decimal points. That was two, so I need to have two numbers behind the decimal point in my answer."	
(B) "I split the problem into two problems to make it easier. So I did $.2 \times .3$ and that got me 0.06. Then I added 0.3 to that and got 0.36."	
(C) "I like to change the problem so that I can use a whole number. I changed this problem to 3×0.12 because I can just multiply the 0.3 by ten, but I have to divide the other number by ten so I don't change the answer."	
(D) "I just multiply like they are fractions. So it's like multiplying $12/10$ and $3/10$. I multiply the 12×3 and the 10×10 and get $36/100$. That's 0.36 when I write it as a decimal."	
(E) "I need to make the length of the numbers the same, so I can line them up. My new problem is $1.2 \times .30$. I multiply them like regular numbers, then I just bring down the decimal point from the .30, so there are two numbers behind the decimal point."	
Mathematical Structure Sample Item	
Ms. Fischer is working with her students on fraction division using a measurement (or quotitive) interpretation, meaning that the quotient specifies the number of equal groups. She wants to give them word problems that use this interpretation of division so that students can practice giving explanations of fraction division.	
Write a word problem that uses a measurement interpretation of division and could be solved using the problem $13 \div \frac{1}{2}$.	

Reflecting on the Framework: Affordances and Constraints

In constructing this framework for the mathematical work of teaching, we sought to develop a tool that could support a focus on the mathematical work that teachers do in the context of every day practice. In reflecting on the framework in its current version, we believe that it does provide a mapping of a practice-based view of contexts in which teachers need to draw on specialized content knowledge. Furthermore, by focusing on a wide array of mathematical objects with which teachers interact, the framework provides insight into the diverse terrain of teachers' knowledge use in teaching. By focusing on the *mathematical* work of teaching, this framework attempts to push the envelope on the types of items that could be written in ways that may not emerge when approaching item writing by starting with the knowledge to be assessed. This framework also provides key insights into the interactions between the mathematical work of teaching and the mathematical content of elementary school in ways that could support item writers to develop assessment tasks within and across different mathematical topics. Despite the potential affordances of the mathematical work of teaching framework, we also recognize that the framework, as a tool for item writing, does not necessarily

provide all of the support that assessment developers might need to write items to measure teachers' specialized content knowledge. In the following section, we describe a set of additional supports that we hypothesize might be needed for item development.

The focus of the framework on the mathematical work of teaching with the additional supports of highlighting interactions with mathematical content may not provide sufficient support for item development if the item writers do not have well-developed mathematical knowledge for teaching across mathematical topics and with respect to actions on different mathematical objects. For example, to write productive items about mathematical structure, item writers would need to know common mathematical structures relevant to K-6 mathematics, such as different interpretations of subtraction or division (i.e., take-away vs. comparison, partitive vs. measurement) or common different problem structures in early addition and subtraction tasks (i.e. result unknown, change unknown). Similarly, in order to write items about the validity and generalizability of student strategies, item developers would need to know common strategies used by children in that content area, such as knowing different valid (or invalid) strategies for comparing fractions (e.g., McNamara & Shaughnessy, 2010). For items about representations as objects, item writers would need to know relevant representations for a particular content area (e.g., area models, number lines, sets, and fraction bars for fractions concepts) and how the key ideas of that content are highlighted or not in different representations, such as knowing how different mathematical ideas of decimal multiplication and place value play out in an area model (Ball, Lubienski, & Mewborn, 2001). To develop items about interpreting student errors, item writers would need to know what are likely errors within particular content areas and what the reasoning behind the errors might be, such as knowing what key mathematical ideas of place value are violated when students incorrectly regroup across zero in multi-digit subtraction (Ball et al., 2008; Fuson, 1990). All of these examples highlight the need for item writers themselves to have well-developed mathematical knowledge for teaching or the tools to access and learn this knowledge themselves in order to develop assessment tasks.

Support for item writers with respect to mathematical knowledge for teaching may be particularly needed for developing items at the lower elementary grades. Much of the key mathematics of those grades is so tacit for adults that they are likely to struggle to determine what ideas could be addressed in items, such as knowing the different mathematical ideas that must be coordinated by children when counting an ill-structured set of objects to determine “how many”, such as one-to-one correspondence, verbal counting, cardinality principle, and strategies for keeping track of what has been counted (Clements & Sarama, 2014; Richardson, 2012). Similarly, to develop items about early place value, item writers would need to know the key, but often tacit, ideas of place value that are not related to operations, such as the role of zero as a place holder or that quantities are represented symbolically left to right. Item writers might also need support in knowing how to construct item scenarios with reasonable approximations of student work at the relevant grade. For example, what would a student's drawing look like when trying to represent fractions with area models? What are reasonable student explanations of their thinking around particular content? Adults with less experience in K-6 classrooms struggle to construct student talk and written work that is reasonably authentic, as adult's own ways of thinking, talking, and representing mathematics are likely to be much more sophisticated than those of children, especially at the lower elementary grades.

Another area in which the mathematical work of teaching, as written, might not be sufficient to support item development around specialized content knowledge is related to the

distinction between what counts as common content knowledge and specialized content knowledge for teaching. In our initial work with this framework, we have found that some sets of actions on mathematical objects seem to sit more clearly in the space of specialized content knowledge, such as analyzing the validity and generalizability of student non-standard strategies, which is in alignment with what was found by Hill, Schilling, and Ball (2004). Other sets of actions emerge at times closer to the line between common and specialized content knowledge, such as making a conjecture. One could argue that this is work that students also often do in mathematics classrooms, but it is important work that is not done in other fields. The line between SCK and CCK maybe particularly challenging to distinguish in the context of the Common Core, since the Standards for Mathematical Practice now ask students (and teachers) to engage with content through mathematical practices such as constructing arguments and critiquing the thinking of others, actions which in some ways align with some of the mathematical work of teaching, such as analyzing the strategies used by others. The intent and nuance of the work may be different when students critique the explanations of peers in a classroom and when teachers are making sense of those strategies but there are some interesting similarities. This points to the possibility that item developers might need further support in how to write items that focus more squarely on specialized content knowledge.

These hypothesized supports serve as an initial set that address some key areas of concern when supporting the development of assessments. There may be other additional challenges that would arise when using the framework for this and other purposes that could require different types of support.

Discussion and Conclusion

In this paper, we have presented a framework to support the development of assessments of mathematical knowledge for teaching (Ball et al., 2008). To consider the theoretical and practical implications of this framework, we first acknowledge that this framework offers one decomposition of the mathematical work of teaching; there may be other useful ways to parse teachers' mathematical work that would foreground different aspects of practice and knowledge use. Furthermore, this framework was developed based on the concept of mathematical tasks, or the mathematical work of teaching (Ball & Bass, 2002; Ball et al., 2008) which were conceptualized based on work in elementary mathematics and with the intended goal of supporting assessment development around elementary MKT. This raises the question of whether the mathematical work of teaching framework proposed here would apply equally well for the work of teaching secondary mathematics or whether there may need to be revisions or additions. For example, at the elementary level, we chose to group explanations and justifications together as an object given the nature of mathematical arguments typically constructed at the elementary level. At the secondary level, it might be more appropriate to include "justification and proof" as a separate mathematical object with which secondary mathematics teachers interact. As part of our future work, we will be pursuing this line of inquiry as we work to support assessment development around secondary MKT.

Despite these potential limitations, the mathematical work of teaching framework offers a contribution that has both theoretical and practical implications. First, this framework builds on and expands upon the mathematical tasks of teaching (Ball & Bass, 2002; Ball et al., 2008) to provide a comprehensive and nuanced identification of the mathematical work that teachers do in the context of teaching. This provides a detailed and practice-based lens (Ball & Bass, 2002) for the contexts when teachers must draw on mathematical knowledge for teaching in their practice.

A novel contribution of this framework is the idea of organizing the work around “mathematical objects” that teachers encounter and interact with in practice. This organization highlights the central role of mathematics in the framework and also affords seeing the diverse ways that teachers interact with different types of mathematical objects (e.g., representations, explanations, mathematical structure). Drawing on a practice-based perspective on mathematical knowledge for teaching, the framework offers a systematic way to identify and examine MKT by focusing first on teaching practice and the nature of teachers’ mathematical work and then including the knowledge needed to manage that work. This perspective is different from starting with teachers’ mathematical knowledge. Our approach explicitly highlights the use of knowledge in practice.

The MWT framework could also serve a number of practical purposes in both assessment development and teacher education. First, the framework was designed with the purpose of supporting the development of assessments of mathematical knowledge for teaching at the elementary level. This tool, along with additional supports described in the previous section, can provide item writers ways to develop assessments of MKT at a larger scale than has previously been possible, when items such as those developed by the Learning Mathematics for Teaching project (Hill, Schilling, & Ball, 2004) have been crafted by groups of experts involving mathematicians, mathematics educators, and teachers. Specifically, this framework could serve as a tool for developing items that appraise mathematical knowledge for teaching in the context of how the knowledge might be used in practice. Furthermore, the MWT framework could provide support in maintaining the focus on the *mathematics* in the work of teaching in ways that help item writers avoid developing items that assess teachers’ pedagogical choices and decisions. Similarly, the framework could help illuminate the specialized knowledge that teachers draw on in their work to support item writers in writing items that assess more than common content knowledge of particular topics. Finally, the MWT framework offers systematic ways for assessment developers to manage the connections between the mathematical work of teaching and mathematical content in ways that would allow for building assessments that tap into the diverse specialized content knowledge needed to teach the K-6 curriculum. As we work with item writers from various backgrounds to develop a content knowledge for teaching assessment for elementary mathematics, we are able to examine the utility and limitations of the MWT framework for supporting assessment development at scale.

Although the mathematical work of teaching framework was developed for the purposes of building assessments, we believe that the framework has the potential to be used for other purposes related to teacher education and professional development, because it offers a systematic identification of the mathematical work of teaching. For example, this framework could be used as an organizing principle for designing mathematics content courses for pre-service teachers to support a focus on the ways that mathematical knowledge is used in practice. Similarly, the framework could be used for as a tool for curricular mapping in mathematics teacher education so that programs could systematically design learning experiences for complementary parts of the framework in different courses (e.g., content vs. methods courses, content courses in different topics such as number or algebra). This framework could also be a useful tool for increasing the MKT of teachers and faculty, including supporting professors and instructors of mathematics content for teachers courses (who likely did not teach elementary school themselves) in better understanding the ways elementary teachers need to use mathematics in practice and the nature of this knowledge.

In the context of the increased mathematical demands of the Common Core State Standards for Mathematics and data showing that many U.S. elementary school teachers lack strong MKT (e.g., Hill & Ball, 2004; Ma, 1999; Tatto et al., 2008), there is an urgent need to develop elementary teachers' mathematical knowledge for teaching. These mathematical demands on teachers are not new (Ball, Hill, & Bass, 2005) and are likely to continue with goal of preparing skillful and responsive practitioners (Ball & Forzani, 2011) Along with ways to support teacher knowledge development, the field needs assessment tools that will allow us to measure and track teachers' growth. The mathematical work of teaching framework contributes a tool to aid in these efforts, especially in working at scale.

Author Note

This research was supported in part by a grant from Bill & Melinda Gates Foundation, entitled "Building Practical Infrastructure for Learning to Teach the Common Core." The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views of the Bill & Melinda Gates Foundation.

Correspondence concerning this article should be addressed to Sarah Kate Selling, 1005D, School of Education Building, University of Michigan, Ann Arbor, MI 48109. Contact: sselling@umich.edu.

References

- Adler, J., & Venkat, H. (2014). Mathematical knowledge for teaching. In S. Lerman (Ed.), *Encyclopedia of Mathematics Education* (pp. 385–388). Springer Netherlands.
- Ball, D. L., & Bass, H. (2002). Toward a practice-based theory of mathematical knowledge for teaching. In B. Davis & E. Simmt (Eds.), *Proceedings of the 2002 annual meeting of the Canadian Mathematics Education Study Group* (pp. 3–14). Edmonton, AB: CMESG/GDEDM.
- Ball, D. L. & Bass, H. (2009). With an eye on the mathematical horizon: Knowing mathematics for teaching to learners' mathematical futures. Paper prepared based on keynote address at the 43rd Jahrestagung für Didaktik der Mathematik held in Oldenburg, Germany, March 1 – 4, 2009.
- Ball, D. L., Hill, H.C, & Bass, H. (2005). Knowing mathematics for teaching: Who knows mathematics well enough to teach third grade, and how can we decide? *American Educator*, 29(1), p. 14–17, 20–22, 43–46.
- Ball, D. L., Lubienski, S. T., & Mewborn, D. S. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. In V. Richardson (Ed.), *Handbook of research on teaching (4th edition)* (pp. 433–456). New York: Macmillan.
- Ball, D. L., & Forzani, F. M. (2011). Building a common core for learning to teach: And connecting professional learning to practice. *American Educator*, 35(2), 17–21, 38–39.
- Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, 59(5), 389–407.

- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., ... & Tsai, Y. M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, 47(1), 133–180.
- Clements, D. H., & Sarama, J. (2014). *Learning and teaching early math: The learning trajectories approach*. New York: Routledge.
- Delaney S. F., Sleep L., Ball D. L., Bass H., Hill H. C. and Dean C. (2005, April). Validating “specialized mathematics knowledge for teaching”: Evidence from mathematicians. Paper presented at the Annual Meeting of the American Educational Research Association, Montréal, Canada.
- Döhrmann, M., Kaiser, G., & Blömeke, S. (2014). The conceptualisation of mathematics competencies in the international teacher education study TEDS-M. In S. Blömeke et al. (Eds.), *International Perspectives on Teacher Knowledge, Beliefs and Opportunities to Learn* (pp. 431–456). Springer Netherlands.
- Eisenhart, M., Borko, H., Underhill, R., Brown, C., Jones, D., & Agard, P. (1993). Conceptual knowledge falls through the cracks: Complexities of learning to teach mathematics for understanding. *Journal for Research in Mathematics Education*, 24(1), 8–40.
- Fuson, K. C. (1990). Issues in place-value and multidigit addition and subtraction learning and teaching. *Journal for Research in Mathematics Education*, 21(4), 273–280.
- 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., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction*, 26(4), 430–511.
- Hill, H. C., Dean, C., & Goffney, I. M. (2007). Assessing elemental and structural validity: Data from teachers, non-teachers, and mathematicians. *Measurement*, 5(2–3), 81–92.
- Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, 42(2), 371–406.
- Hill, H. C., Schilling, S. G., & Ball, D. L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *The Elementary School Journal*, 105(1), 11–30.
- Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Mahwah, NJ: Lawrence Erlbaum Associates.
- McNamara, J., & Shaughnessy, M. M. (2010). *Beyond pizzas & pies: 10 essential strategies for supporting fraction sense, grades 3–5*. Math Solutions.
- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for Mathematics*. Washington, DC: Authors.

- Richardson, K. (2012). *How children learn number concepts: A guide to the critical learning phases*. Bellingham: Math Perspectives Teacher Development Center.
- Rowland, T., Huckstep, P., & Thwaites, A. (2005). Elementary teachers' mathematics subject knowledge: The knowledge quartet and the case of Naomi. *Journal of Mathematics Teacher Education*, 8(3), 255–281.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2) 4–14.
- Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1–23
- Tatto, M. T., Ingvarson, L., Schwille, J., Peck, R., Senk, S. L., & Rowley, G. (2008). *Teacher education and development study in mathematics (TEDS-M): Policy, practice, and readiness to teach primary and secondary mathematics*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.
- Wilson, S. M., Shulman, L. S., & Richert, A. E. (1987). “150 different ways” of knowing: Representations of knowledge in teaching. In J. Calderhead (Ed.), *Exploring teachers' thinking* (pp. 104–124). London: Cassell

Assessing Mathematical Knowledge for Teaching: The Role of Teaching Context

Geoffrey Phelps
Educational Testing Service

Heather Howell
Educational Testing Service

Abstract: Assessments of mathematical knowledge for teaching (MKT), which are often designed to measure specialized types of mathematical knowledge, typically include a representation of teaching practice in the assessment task. This analysis makes use of an existing, validated set of 10 assessment tasks to both describe and explore the function of the teaching contexts represented. We found that teaching context serves a variety of functions, some more critical than others. These context features play an important role in both the design of assessments of MKT and the types of mathematical knowledge assessed.

Keywords: teacher content knowledge, mathematical knowledge for teaching, teacher assessment, mathematics, assessment

Introduction

Mathematical Knowledge for Teaching (MKT) is the content knowledge used in recognizing, understanding, and responding to the mathematical problems and tasks encountered in teaching the subject (Ball & Bass, 2002; Ball, Thames & Phelps, 2008). Assessments of MKT are designed to measure the mathematical knowledge that teachers use in these teaching practices. A number of practice-based assessments of MKT have recently been developed for teachers of K-12 grades (Herbst & Kosko, 2014; Hill, Ball, & Schilling, 2008; Hill, Schilling, & Ball, 2004; Kersting, 2008; Krauss, Baumert, & Blum, 2008; McCrory, Floden, Ferrini-Mundy, Reckase, & Senk, 2012; Phelps, Weren, Croft, & Gitomer, 2014; Tatto et al., 2008).

We follow Ball, Thames, and Phelps (2008) in defining MKT to include the full range of mathematics content knowledge used in teaching. The most widely assessed component of MKT is the common content knowledge that is taught and learned as part of regular schooling and is familiar to most adults. There is a long history of assessing teachers' common mathematical knowledge (Hill, Sleep, Lewis, & Ball, 2007). Often these assessment tasks look identical to those on student assessments because the construct is essentially the content of the student curriculum, either at grade level or at a level above the assigned grade (Phelps, Howell, & Kirui, 2015).

MKT assessments have generally focused, however, on the specialized forms of content knowledge that only teachers need to use in the course of their day-to-day work (Ball et al., 2008). While definitions and focus vary in the literature, and the mapping of the MKT

construct is likely somewhat dependent on curriculum and culture, most studies share a focus on MKT as a form of applied knowledge that goes beyond common content knowledge (Krauss et al., 2008; McCrory et al., 2012; Thompson, 2015; Turner & Rowland, 2008). MKT assessments typically present teachers with content tasks that are encountered in teaching, such as interpreting student thinking and work, selecting materials for instruction, explaining concepts and procedures, or evaluating whether to use a representation for a particular instructional purpose (Ball & Bass, 2002; Hill et al., 2004). And since these tasks often occur in complex instructional contexts, MKT assessments typically also provide key information about the teaching context, such as the learning goals that direct the teaching, details about a student's prior academic work, or how students are grouped and organized (Phelps et al., 2015). Assessments of MKT differ in how teaching practice is represented. Some provide written descriptions, while others incorporate video or animations depicting mathematics teaching (see, for example, Herbst & Kosko, 2014; Hill et al., 2004; Kersting, 2008). These features of context support test takers in recognizing the relevant aspects of the content task, understanding the content problem, or providing a response to the assessment question.

This contextualization of MKT assessment tasks is in part theoretically motivated. Ball and Bass (2002) argue that *how* teachers encounter mathematics in their teaching directly shapes the nature of the mathematical knowledge that is needed. The context used in many MKT assessment tasks defines both *what* kinds of content knowledge teachers need to use and *how* they use this knowledge. Largely missing, however, from the current literature on MKT assessment are well-articulated design arguments that make clear the links between the construct and assessment task design (Mislevy & Haertel, 2006). Given the central role of teaching in MKT, it seems likely that any endeavor to assess MKT would require consideration of how context functions in the design of MKT tasks (Phelps et al., 2015).

In this study, we take the first steps in this direction by presenting arguments and illustrations for how context functions in a set of elementary-level MKT assessment tasks, with a particular focus on how context enables tasks to measure MKT that goes beyond common content knowledge. We do not take up the question of whether other sub-components of MKT are distinctly measureable, as other studies have done (see, for example, Hill et al. (2008) and Krauss et al. (2008) for different approaches to the measurement of PCK as a distinct domain). Our argument is simply that context matters in the assessment of some components of MKT more than others; in particular it matters more for components that go beyond common content knowledge. Because these types of knowledge have been the objects of intense interest in teacher education it is worth attending closely to how context matters in their assessment.

The paper is organized as follows. First, we discuss the role of context in establishing the construct validity of MKT assessments using illustrative examples. We follow Messick's (1989) view of construct validity, which helps to determine how relevant and representative the tasks are in measuring MKT. We begin with an example that includes three tasks that assess similar content focused on exponential expressions but vary in how teaching context is represented in the task. This set of tasks provides a concrete illustration of major differences in context and its function. Next, we discuss two task examples in detail to illustrate the design and content focus of MKT assessment tasks and to make clear our arguments about the role that context plays in these assessment tasks. Finally we present a summary of how context functions across the 10 tasks and discuss the implications for assessing MKT.

The Role of Teaching Context in Assessing MKT

The appropriate use of teaching context in the assessment of MKT can help avoid threats to construct validity, namely construct-irrelevant variance and construct under-representation. Construct-irrelevant variance occurs when an assessment represents dimensions that are irrelevant to the correct interpretation of the construct, and construct under-representation occurs when an assessment does not adequately represent the dimensions of the construct that are the focus of the assessment (Messick, 1989). In respect to MKT, many assessments are designed to measure the MKT that is specialized to the work of teaching. In cases where teaching context is critical to assessing particular aspects of MKT, the absence of teaching context could lead to construct under-representation.

We begin with an illustration designed to highlight the various roles that context can play in the measurement of MKT. We present three related example tasks in Figure 1. The example in panel C was developed for the Measures of Effective Teaching project (Phelps et al., 2014) and is one of the 10 tasks analyzed in this study. Task selection and analysis is addressed in more detail in the methods section. The examples in panels A and B of Figure 1 are variants created by the authors for illustrative purposes to demonstrate both when teaching context does and does not support the assessment of MKT.

A. Common Content Knowledge	B. Common Content Knowledge in a Teaching Context	C. Specialized Content Knowledge in a Teaching Context												
<p>Evaluate each of the following simple exponential expressions.</p> <p>$3^3 =$</p> <p>$2^3 =$</p> <p>$2^2 =$</p>	<p>Ms. Hupman is teaching an introductory lesson on exponents. She gives her students a set of problems to check their proficiency in evaluating simple exponential expressions. Ms. Hupman looks over the work from one of her students. For each of the answers, indicate if the student's evaluation is correct or incorrect.</p> <table> <tr> <th></th><th>Correct</th><th>Incorrect</th></tr> <tr> <td>$3^3 = 9$</td><td>☐</td><td>☐</td></tr> <tr> <td>$2^3 = 6$</td><td>☐</td><td>☐</td></tr> <tr> <td>$2^2 = 4$</td><td>☐</td><td>☐</td></tr> </table>		Correct	Incorrect	$3^3 = 9$	☐	☐	$2^3 = 6$	☐	☐	$2^2 = 4$	☐	☐	<p>Ms. Hupman is teaching an introductory lesson on exponents. She wants to give her students a quick problem at the end of class to check their proficiency in evaluating simple exponential expressions. Of the following expressions, which would be <u>least</u> useful in assessing student proficiency in evaluating simple exponential expressions?</p> <p>☐ 3^3</p> <p>☐ 2^3</p> <p>☐ 2^2</p>
	Correct	Incorrect												
$3^3 = 9$	☐	☐												
$2^3 = 6$	☐	☐												
$2^2 = 4$	☐	☐												
Key: 27, 8, 4	Key: incorrect, incorrect, correct	Key: 2^2												

Figure 1. Tasks to illustrate differences in types of content knowledge assessment.

Each of these three tasks involves the same underlying mathematical content, but they differ in whether and how each is situated in teaching. Task A does not include a context and simply requires the test taker to evaluate three exponential expressions. This task is not situated in teaching other than representing mathematics that is part of the grade school curriculum. However, the absence of context in this task is construct relevant because no context is required to assess whether teachers can do the work of the student curriculum.

Task B includes a context that shows a student's evaluation of three simple exponential expressions. The student has answered two problems incorrectly and one correctly. The test taker does not need to draw conclusions about *why* the student answered each correctly or incorrectly. He only needs to evaluate each problem and check the correct answer against the student's answer to determine whether the student's answer is correct or incorrect. While the look and feel of Tasks A and B are different, the mathematical work and knowledge required to answer is essentially the same. Both measure a test taker's ability to evaluate expressions. The context in Task B is arguably *construct-irrelevant* (Messick, 1989), meaning that its presence or absence does not relate directly to the skill of exponent arithmetic. However, the longer text included in Task B increases the reading burden on the test taker, raising the possibility that the task might unintentionally measure reading ability in addition to the skill of exponent arithmetic. Reading load is not necessarily problematic; the text is not excessive in length and the level of reading required may be well within the abilities of the tested population. But to the extent that such a task measures something unintended (in this case, reading ability), it can be a source of *construct-irrelevant variance* in the test scores (Messick, 1989).

Task C, like Task B, includes a written teaching scenario. But in this case, the context serves to direct the test taker to consider which expression would be a poor choice for teachers to use in understanding whether students know how to evaluate expressions. To respond to this task, the test taker needs to already know, or know how to figure out, what kinds of confusion students are likely to exhibit (e.g., confusion about which number is the base or exponent or confusion around what kind of operation is required to evaluate the expression). The test taker then needs to anticipate what the solution to each of the problems would be using the incorrect methods students might apply and from this figure out which problems reveal these confusions. The mathematical knowledge involved in responding to this task goes beyond the common content knowledge of how to evaluate exponents. The context that is included in this task is relatively minimal but clearly necessary; without the context the test taker lacks key information for comparing the problem choices. Unlike Task B, where the context is *irrelevant* to the content assessed, in Task C the context is *relevant* and arguably critical to the content knowledge that is being assessed.

The three tasks shown in Figure 1 are intended to illustrate a number of key points in the design of MKT assessment tasks. First, context is not always needed. Most notably, as illustrated in panel A, when teachers are simply doing the math that their students are learning, there is likely no need for context (Phelps, Howell, Schilling, & Liu, 2015). The context in Task B illustrates an authentic situation in teaching that requires the teacher to have common content knowledge. But from an assessment perspective, when measuring this type of MKT, it will often be more efficient to present the task without a context, as illustrated in Task A. A basic principle of assessment design is that irrelevant context should be avoided to the greatest extent possible so that only the intended construct is measured (Messick, 1989).

Second, as illustrated by task C, context can play a critical and relevant role in assessing the construct when the goal is to assess the components of MKT that go beyond the mathematics that students are expected to master (i.e., SCK or PCK in the Ball et al. (2008) model). In such cases, eliminating the context might shift the focus of the task in ways that leave the test taker unsure what is being asked or might fundamentally change the content assessed. Eliminating context entirely could reduce tests of MKT to assessing only the types of common content knowledge illustrated by task A, which would lead to tests that suffered from threats of construct under-representation (Messick, 1989).

Figure 1 also illustrates that it is not always simple to determine whether context is relevant. At first glance, Tasks B and C seem quite similar. It is only through analysis of the work that each task requires of the test taker and consideration of the measured construct that such a determination can be made. Consequently, from an assessment design perspective, it is critical to clarify how context that is included in an assessment task is relevant to the particular features of the construct being assessed (Mislevy & Haertel, 2006).

Methods

Our goal in this study was to systematically investigate the ways in which teaching context can function in tasks designed to elicit the types of MKT that are particular to the work of teaching mathematics. While we follow general procedures for qualitative coding, our method differs from typical qualitative work in two key ways. First, our ‘data’ are the tasks themselves. We selected a set of tasks for which we have a large set of ancillary data showing that they perform well as measures of CKT and that context matters in how respondents reason through each task. We did not, however, examine teachers’ actual response data in this particular study. Our claims therefore are built on arguments about task design and not on empirical data comprised of test takers’ responses. Therefore, our results are the categories and associated characteristics of task design that emerged in the course of the close analysis of the MKT tasks. We think this type of close, rigorous analysis helps to call attention to aspects of task design that are otherwise largely invisible, even to test designers. We describe the process in some detail to help the reader follow our logic.

Selection of MKT Tasks for Analysis

The analysis that follows focuses on a set of 10 mathematics tasks that were developed as part of the Measures of Effective Teaching project to measure elementary level MKT (Phelps et al., 2014). These tasks were chosen because we had strong evidence from a prior cognitive interview study that they situated test takers in teaching practice as designed (Gitomer, Phelps, Weren, Howell, & Croft, 2014; Howell, Phelps, Croft, Kirui, & Gitomer, 2013). As part of that study, we wrote rationales detailing the embedded assumptions about how context would function and about the construct each task measured. The study established that the alignment of participant reasoning to these rationales was strongly related to answering correctly or incorrectly. Across all mathematics task level interview responses ($n = 640$), 88% showed the desired pattern in which correct answers matched pre-specified correct reasoning and incorrect answers did not match that correct reasoning. For 97% of responses, participants reported that the task was an authentic representation of actual teaching practice. The study also found no evidence that reading load introduced construct-irrelevant variance by interfering with test takers’ interaction with the assessment tasks (Gitomer et al., 2014; Howell et al., 2013). These response patterns led us to conclude that knowledgeable teachers were situated in context as specified by the task design.

It is worth clarifying that our goal was not to generalize to all MKT tasks or other such practice-based items. Instead we used strong tasks from a prior study with the goal of using this selection as a site for naming and defining important task design characteristics. Specifically, in order to understand how context can function, we required a set of tasks that measure more than common content knowledge, in which context is available to be analyzed, and for which we have some evidence that the context serves a function.

Analytic Method

As a first step in the analysis we expanded the written rationales used in the prior study to account more explicitly for context features and to understand better the role that context played in these tasks (Howell et al., 2013). We started by simply describing the context and its role in shaping how the test taker interacts with the content problem. These descriptions constituted the first step in our qualitative analysis and subsequently became objects of the second step of analysis. A summary of such a description is provided below for the task shown in figure 2.

To assess her students' prior knowledge about evaluating arithmetic expressions, Ms. Santiago assigned a worksheet of problems. She noticed that Alexis answered the first two incorrectly and the next two correctly.

$$\begin{array}{ll} 1) 7 \times 2 - 6 + 3 & = 5 \\ 2) 9 - 5 + (16 \div 8) & = 2 \\ 3) 9 + 24 \div 3 - 1 & = 16 \\ 4) 17 - (3 + 7 \times 2) & = 0 \end{array}$$

Which of the remaining problems is Alexis likely to answer incorrectly?

- ☐ $8 + 7 - 12 \div 3$
- ☐ $13 - 3 \times 2 + 5$
- ☐ $(27 \div 3 - 4) + 8$
- ☐ $(16 - 12) \times 5 + 10$

Figure 2. The Santiago task.

To respond to this task, the test taker needs to analyze the four examples of Alexis' work, determine what she did to get the first two problems wrong, and then test any hypotheses about her confusion to see if they are consistent with answering the other problems correctly. The test taker needs to select an option that Alexis would answer incorrectly, assuming Alexis persists in the same error. However, the underlying, important task is to figure out what Alexis is misunderstanding. The assessment task is focused on the

recurrent teaching practice of diagnosing student understandings or misunderstandings based on the written work they produce.

Analysis of the given problems reveals that in each of the incorrect problems, Alexis has added before subtracting. In the first problem she added 6 and 3 first and then subtracted the total of 9 rather than subtracting 6 and then adding 3. In the second problem she added 5 and 2 (where 2 is the result of 16 divided by 8) and then subtracted the total of 7 rather than subtracting 5 then adding 2. However, in the third and fourth problems this particular error does not lead to an incorrect answer. In the third problem, the ordering of the operations happens to be such that adding before subtracting is appropriate. In the fourth problem, the parentheses indicate that the expression inside should be added first before subtracting. There is not enough evidence to know *why* Alexis is making this error, although experienced teachers may recognize it as a possible overgeneralization of the use of the mnemonic PEMDAS¹ to dictate the order of operations. If we assume that Alexis will persist in the same error, the second answer option is the only option she would answer incorrectly because for each of the others, like the third and fourth given problems, adding before subtracting happens to be correct.

The scenario only specifies “arithmetic expressions” as the content topic under study, but the form in which the mathematics problems are written provides a great deal of subtle contextual information about the level of the students. Each expression is written out as a single line, using the division symbol $\div$ and the multiplication symbol $\times$ rather than a fraction bar for division or a dot for multiplication. All four operations (addition, subtraction, multiplication, and division) are represented and parentheses are used, but there are no exponents. These details communicate to someone with knowledge about the teaching of this mathematics that the students are likely studying order of operations. Their use of the operations themselves is likely fluent at this point, but their ability to combine the operations correctly may not be. In the context of the assessment task, this is important because it makes some possible errors far less likely. For example, one could have assumed that Alexis misread the addition symbol or that she did not know how to perform the subtraction correctly, but this is an unlikely error for a student who is working with expressions of this type.

On the other hand, it is quite common for students at this level to make mistakes in the ordering of the operations. While the scenario does not state that this is an order of operations problem, the contextual clues embedded in the format of the content problems themselves make the work the test taker needs to do much easier by narrowing the field of all possible errors to a fairly small set of likely ones that need to be considered. This is a critical piece of information because it allows the test taker to rule out other competing, but unlikely theories. Again, one reason this set of assessment tasks was useful to study is that the prior interview work provides evidence to support such claims about the functioning of the context. And indeed in a prior study using this task, participants often referred explicitly to it being about order of operations, confirming this part of the design theory (Howell et al., 2013).

¹ PEMDAS is a mnemonic device commonly used in the U.S. to help students remember the order of operations. It stands for “parenthesis, exponents, multiplication, division, addition, subtraction,” and is not strictly mathematically correct as written, although when used in instruction teachers generally qualify it by stating that the pairs “MD” and “AS” are performed in order, left to right, at the same time, not one before the other as the device implies.

The context also includes information about the student, Alexis, stating that she answered the first two problems incorrectly and the second two correctly. It is not strictly necessary to state which are correct and which incorrect, but providing the information up front may decrease the cognitive load on the test taker and encourage him to focus on the student's thinking rather than on whether the problems are correct. And pointing out that these are Alexis's answers also conveys a crucial piece of information about what the test taker needs to do by setting the condition to be met—the identified misconception must explain both Alexis's correct and incorrect work, and it must be a systematic error that the student makes consistently. A test taker who fails to attend to this aspect of the context may read through the problems assigning a unique diagnosis to each, or may cite difficulties students generally have with such problems without determining the specific difficulty Alexis is having. Both were patterns we observed in prior interview data and were associated with incorrect answers (Howell et al., 2013).

Finally, the assessment task presents an authentic scenario. Teachers frequently have to draw conclusions about student thinking from written work. The task of figuring out what Alexis is thinking seems not just plausible but worthwhile; teachers can't make informed decisions about next instructional steps without knowing first what their students understand and do not understand.

The summary above illustrates the type of descriptive account that was generated for each of the 10 tasks. These accounts provided rich descriptions of how the tasks functioned and more specifically the role that context played in these tasks. They also were used as the basis for generating provisional statements describing each context element. We then coded each identified context element inductively with short phrases describing the ways in which the context element functioned in the test taker's anticipated interaction with the task. We used a constant comparative method (Strauss & Corbin, 1990) to do this coding, which can be described in four steps: (1) independently analyzing a subset of tasks, (2) reconciling the coded elements and functions across tasks, (3) revising the list to reflect all elements and functions and testing the new categories by recoding the subset of tasks, and (4) expanding and iterating to a larger set of tasks until we had reached consensus on all codes for all context elements observed across all 10 tasks. Our goal in this work was not to achieve a particular level of coding reliability, but rather to generate a useful set of categories that captured the types of elements and functions we saw both in a given task and collectively across tasks. The short descriptors of the functions were then grouped together to form more general categories, and the entire set of tasks reviewed and recoded using these categories.

This process of task analysis generated three sets of categories that were relevant to describing the context and its function. Because we view these categories as an important outcome of this study, they are described in more detail below in the results section.

Results

The results are organized in two main sections. The first section presents the categories that were derived inductively from the analysis of the 10 MKT tasks. The second section focuses on the use of these categories to describe the context features and their function across these 10 tasks. While we present counts across the set of tasks to illustrate the frequency, distribution, and co-occurrences we observed, we remind the reader that for a study of this type the main results are the identification and description of the categories themselves.

Teaching Context and Function

Context focus. The various teaching contexts identified in the MKT tasks mapped onto three major components of instruction. These included features of *students* such as their history, learning needs, and actions; the *content* and how it is situated in the curriculum of school learning; and, the *setting*, which includes class size or grouping or mode of instruction such as lecture or discussion. Not only are these particular features central to instruction, but they have also recurred in many different heuristics and models used to characterize instruction (see, for example, Cohen, Raudenbush, & Ball, 2003; Hawkins, 1974; McDonald, 1992; Schwab, 1978). For each of the 10 MKT tasks, elements of the context could be identified as providing context for the *content*, *student*, or *setting* of instruction. These categories are useful for identifying the aspects of instruction that are the focus of the context features.

Context Function. The categories that were derived from the analysis describe the main function of the contexts identified in the 10 tasks. These categories are described in Table 1.

Table 1. *Context Functions.*

	Context function	Description
Critical functions	Narrows a set of possibilities	Context that functions by narrowing the set of possibilities that the test taker must consider – e.g., narrowing the possible answer choices or eliminating one or more options. This can be quite subtle, as in cases where the specified level of a student or class sets an expectation for the level of sophistication one might expect in an answer, which in turn serves to eliminate some set of possibilities. Sometimes it might be some other list that is narrowed rather than the answer choices. For example, in the task shared in Figure 2 the test taker needs to figure out what error the student has made before even considering the options, and the content context serves to narrow the possible errors.
	Sets condition for the answer	Context that functions to specify, explicitly or implicitly, what condition the answer needs to meet to be correct. For example, in the task shared in Figure 3, the setting context sets a condition for the answer – i.e., that selected problem needs to be one for which the student's answer will reveal the suspected misconception to the teacher.
Helpful functions	Direct the test taker's focus	Context elements that encourage the test taker to focus (or not to focus) on a particular aspect of the task. For example, in the task shared in Figure 2, the statement that the student answered two problems correctly and two incorrectly is intended in part to cue the test taker to pay attention to the correct work and not just the incorrect work.
	Provides additional information	Context that provides additional information that is useful but not critical. This might include defining a term that some test takers may not know. Or it may include context that reduces cognitive demand by stating up front that a student's work is incorrect so that the test taker knows that figuring this out is not part of the work he needs to do.
	Reinforces critical information	Context that reinforces a key idea. This can help ensure that a test taker is directed to pay attention to critical information and thus raise the likelihood that the test taker engages in the assessment task as intended.
Functions related to face validity	Authenticity	Context that helps support an authentic representation of the work of teaching. Perceived authenticity can be a key motivating factor and enhance validity.
	Plausibility	Context that specifically helps to add plausibility to an element of the task that would not otherwise seem reasonable. For example, in the task shown in panel C of Figure 1, the specification that the problem is a quick check at the end of class makes it feel reasonable that the teacher has a need to diagnose understanding on the basis of a single answer alone. Without this information, the test taker might wonder why the teacher does not simply ask the students to explain their work.
	Motivation	Context that creates a situation in which the test taker can better recognize the importance the task. For example, tasks that give specifics about a student and their learning needs can motivate because there seems to be a real and pressing need to help the student.

Context Relevance. Another pattern that emerged was one of relative levels of relevance or criticality. Some context functions, like those that set the condition the answer needs to meet to be correct, are essential for assessing the MKT content. Without that information the test taker would be unable to respond correctly and the particular type of MKT could not be assessed. Other functions were less essential in that it would still be possible to respond correctly absent that context. However, many of these context features were still quite helpful in directing the test taker and thus might serve to reduce cognitive load. For example, context that functions either to further define a key idea or direct the test taker to pay attention to something important falls into this second group. A third type of context functions to increase face validity, support the test taker's perception of authenticity, or to motivate in other ways that support completing an assessment task.

Teaching Context in MKT Assessment Tasks

To make these three sets of categories more concrete, an example task (Figure 3) is used to illustrate the process and the types of decisions that the coding and classification entailed.

Mr. Chamberlain is concerned that his students' use of the calculator has led them to view the equal sign as a signal to carry out an operation rather than as a symbol indicating equality. Of the following missing-number problems, which would best assess whether students understand the mathematically correct meaning of the equal sign?

- ☐ $__ + __ = 18$
- ☐ $7 + 5 = __ + 6$
- ☐ $__ = 17 + 9 + 5$
- ☐ $23 + 4 = __ = 4 + 23$

Figure 3. The Chamberlain task.

The context for content in this task is given directly and indirectly. The scenario indicates that Mr. Chamberlain's concern is focused on the meaning of the equals sign. The format of the missing number equation problems communicates the level of the students as early elementary and signals that the use of the equal sign is likely new to them. This bolsters the authenticity and appropriateness of Mr. Chamberlain's concern as represented in the problem, as students often misunderstand the equals sign to be a command to perform an operation. It both makes sense that students working at this level would have this confusion and it conveys that the confusion is important for a teacher to attend to. Thus, in this case, the content context provides authenticity and contributes to the face validity of the task.

Unlike the task in Figure 2, in which information about the student was given directly, the student context in this problem is given indirectly in the form of the teacher's concern. What we know about the students is that they have used a calculator, and further that the teacher believes they may hold a particular misconception (that the equals sign is a command

to perform an operation). Knowing that the suspected misconception is connected to calculator use in this way provides key information to the test taker by defining, if indirectly, the operational view of the equals sign. Understanding the difference between the operational and equality views of the equal sign is key to answering correctly, and this piece of context reduces the cognitive load for a test taker unfamiliar with the misconception or with the terminology used to describe it. It also provides a plausible basis for the students to have that misconception, as calculator use is common and can lead to exactly this type of misunderstanding. The student context here serves dual functions. It supports the plausibility of the scenario, contributing additionally to face validity, and it also provides helpful but non-critical information to the test taker by defining a key idea.

We point out here an ambiguity in the coding and classification. One could argue that the teacher's concern is a part of the setting context, and not really information about the students. We acknowledge this, and use this example to draw attention to a necessary imprecision in the categories we have proposed. In many cases the distinctions are subtle and a piece of context might well fall into multiple categories. In fact, in this case we listed the teacher's concern about the operational view as setting context as well as coding the student's use of the calculator as student context. As a feature of the setting, the teacher's concern motivates the task by providing a plausible reason to care which problem is selected, further supporting face validity. More importantly, it sets the condition the answer needs to meet in order to be correct; the correct answer must be a problem that will reveal the given misconception to the teacher. This function of context (setting the condition the answer needs to meet to be correct) is at the highest level of relevance because it is critical that it be included in order for the task to function as designed. That the context is difficult to assign to the categories of setting or student is less important than the critical function it serves in orienting the test taker's thinking. We draw the reader's attention to the ambiguity here to illustrate clearly that our goal is not to create strict divisions between context types so much as to name categories that are useful for systematically analyzing or generating MKT tasks.

We also draw a distinction between the context that is situating the test taker and the actual knowledge or ability that the test taker must have in order to respond to the task correctly. This last piece of context sets the condition the answer needs to meet, and the test taker must distill this understanding from the context in order to answer correctly. But the test taker still needs to know which problem will meet that condition. While the context clues situate the test taker so that she is answering the right question, they do not answer the question for her. In this case, the test taker still needs to know or be able to anticipate that a student with the given misconception will likely write 12 in the blank on the second problem, having interpreted the equal sign as a command to add 5 and 7. For this option, 12 is incorrect because $5 + 7$ is not equal to $12 + 6$. While the student might think about the equal sign incorrectly in each of the other options, the answer the student gives would be the same as the correct answer and would not reveal the error to the teacher. This is the only problem that makes the misconception visible.

Table 2 gives an overview of the context features coded for each of the three MKT tasks that have been discussed in depth so far in the paper (Figures 1, 2, & 3). It is worth noting that while we made efforts to reach consensus in the coding, we do not propose that the context elements for which we coded are fixed or that there is always a clear classification. Rather, we find these elements useful in providing conceptual tools that help to identify and

understand the function of context. Specifically, this makes these context elements more visible and provides a language that can be used to evaluate and critique the design of assessment tasks. The examples in Table 2 also illustrate that not every type of context element or function appears in every task. This is typical of what was represented across the set of analyzed tasks and suggests as a cautionary note that while the proposed categories are analytically useful, they are not strictly necessary. They do not form a template for assessment task construction.

Table 2. *Sample Coding Classifications.*

Task Description	Content context and its function	Student context and its function	Setting context and its function
Ms. Hupman wants to select a brief assessment problem to ascertain whether her students understand how to evaluate exponential expressions. (Figure 1, panel C)	That the lesson is introductory narrows the likely errors students would make¹ to those above the level of arithmetic (students probably know how to multiply).	No student context is given.	That the problem is a quick proficiency check provides plausibility³ for why the answer alone needs to convey information and also sets the condition for the answer¹ – that it must reveal to the teacher whether or not the student is proficient. The focus on the <u>least</u> useful problem decreases authenticity³ as a teacher would generally look for the most useful, not the least.
One of Ms. Santiago's students has answered two order of operations problems correctly and two incorrectly, and the test taker must figure out what she has done wrong and predict which additional problems she will answer incorrectly. (Figure 2)	The types of mathematical symbols used ("x" for multiplication, for example) coupled with the specification that this is prior knowledge for the students narrows the possible error types¹ to exclude arithmetic errors and include errors related to ordering of steps.	The specification that the student answered two problems correctly and two incorrectly encourages the test taker to attend to both² the correct and the incorrect work, suggests a systematic error² , and sets the condition¹ the selected error needs to meet – it needs to explain the given work.	That the student work shown was in response to a worksheet suggests that the teacher is looking at the work after the fact, with time to reflect, making the work needed to analyze the errors more plausible³ .

Mr. Chamberlain is concerned that his students may have a specific misconception about the equal sign and must choose an assessment problem that would reveal to him whether or not they have that misconception. (Figure 3)	The format of the missing number equation problems communicates the level of the students as early elementary and hints that the use of the equal sign symbol is likely new to them, supporting the authenticity ³ of the teacher's concern that they might not understand it.	The teacher expresses concern that students use of the calculator may have caused them to have an operational view of the equals sign, providing a plausible ³ explanation for why they would misunderstand, and defines ² for the test taker what is meant by the operational view.	The teacher's concern that the students may have an operational view of the equals sign provides motivation ³ for the task of teaching, as well as setting the condition ¹ to evaluate the answers as those that reveal that incorrect operational view.
--	--	--	--

Note: Bold text indicates the function of a context element: 1) critical context function, 2) useful context function, 3) face validity context function. A full version of this table and all tasks analyzed is available from the corresponding author upon request.

Table 3 provides an overview of how often each function type for each context category appeared across the 10 tasks analyzed. For many of these tasks, various context functions and features could appear multiple times. For example, the task presented in Figure 3 was coded as having content context that supports authenticity, student context that makes the situation more plausible and defines a key term, and setting context that motivates the situation as well as setting the condition the answer needs to meet. This particular task contributes one count to the content context category and two counts each to student and setting categories.

Table 3. *Context Type and Function for Ten MKT Tasks.*

Type of function	Type of context			Total occurrences over 10 tasks
	<u>Content</u>	<u>Student</u>	<u>Setting</u>	<u>Total</u>
Critical context functions:				
Narrows a set of possibilities	5	2	0	7
Sets the condition for the answer	2	3	2	7
Helpful context functions:				
Directs the test takers focus	1	4	1	6
Provides additional information	4	4	0	8
Reinforces critical information	1	1	0	2
Face Validity				
Authenticity	1	2	2	5
Plausibility	0	1	8	9
Motivation	0	2	1	3

Our summary suggests that for the 10 tasks analyzed the context functions are relatively equally represented across all the major coding categories. While critical functions seem to occur slightly less often in the “setting” column than elsewhere, and face validity functions noticeably more, the overall distribution suggests that all three types of context elements can serve all functions and can also vary in their criticality. This suggests that teaching context, at least as it appears in these particular MKT tasks, can play a variety of functions across a number of major features of instruction.

Discussion

Mathematical knowledge for teaching includes the full range of mathematics used in teaching the subject. This is a form of applied knowledge that teachers draw on and use as they engage in and carry out the many practices that make up the moment-to-moment and day-to-day work of mathematics teaching (Ball & Bass, 2002; Ball et al., 2008). In this study, we conducted an analysis of 10 tasks designed to assess MKT. These tasks assess types of MKT that go beyond the common content knowledge used in doing the work of the student curriculum (e.g., the first example in Figure 1), with the goal of measuring specialized types of MKT used in practices only encountered in mathematics teaching (Phelps et al., 2014). Because these tasks focus on types of MKT applied in teaching practice, they all include teaching context. We found that across these tasks the context served a number of different functions. In fact, for many tasks, the context served multiple functions. We coded almost 50 instances of context serving an identified function across just 10 tasks.

These context features focus on different aspects of instruction. We grouped these under the larger categories of content, student, and setting. These categories provide a useful set of lenses for considering which core aspects of instruction are represented in the context. It also seems likely that different types of MKT tasks might require context that focuses on aspects of instruction that did not come up in our analysis. For example, tasks like the Chamberlain task provide background information about the teacher’s concern. This suggests that *teacher* might be an additional useful category that simply did not appear often in the set of tasks we examined. This category might include information such as teachers’ pedagogical motivations, purposes, or constraints. Other MKT assessment tasks might call for a more fine-grained list of the major components or aspects of instruction, such as separate categories for curriculum materials and content.

We also identified across the context features a variety of functions (Table 1). Again, it is important to emphasize that these functions are almost certainly not exhaustive. Additional functions might be identified for a different set of MKT tasks. Although the list of context functions is likely incomplete, we think it nonetheless provides a useful start and important insights into the assessment of MKT. One insight that emerged from this analysis was that these functions could be placed into three larger groups describing the degree to which the context was critical to assessing the MKT construct. We discuss each group of functions briefly below.

We described one group as “functions related to face validity” (Table 1) because their only role was to support the test taker’s perception of the situation as authentic, to make the work seem plausible, or to motivate. This group of context functions is arguably the least critical for supporting the test taker in providing an answer. In fact, in some situations, these

context features may not be needed at all. If the test taker, for example, is familiar with the content and accepts that it is important and used in teaching, the context may do little more than add to reading load and may even introduce construct irrelevant variance. On the other hand, context that adds face validity can support the test taker in important ways. Michael Kane (2006), in his seminal chapter on validity, argues that tests that lack face validity can introduce construct-irrelevant variance since the test may in part measure a test taker's disengagement with the tasks rather than the construct of interest. Context associated with this group of context functions should be examined with special care to make sure that it plays a sufficiently important role to be included in the assessment task.

We described a second group of context functions as “helpful functions” (Table 1) because they served to support the test taker in providing an answer (e.g., directing the test taker's focus toward a particular aspect of the work or reinforcing critical information). While this type of context was not critical to answering the task, it played an important role, often reducing burden for the test taker. As was the case for the face validity functions, the context associated with this second group is not critical to answering. However, it is not obvious that the context is construct irrelevant, since it appears to support the test taker in productively and efficiently engaging the task.

We described the final group of functions as “critical functions” (Table 1) because the test taker needs to consider the associated context in order to provide an answer. This included cases where the context information narrows the answer possibilities or sets a condition the answer needs to meet. If the context were removed entirely for these instances then the MKT that was the focus of the task simply could not be assessed. In these cases the context is not only critical, but arguably an integral part of the construct itself (Phelps, Howell, & Kirui, 2015). Removing context from these tasks would fundamentally change the MKT assessed and would likely lead to tests that suffered from construct under-representation (Messick, 1989).

Our analysis also revealed that because context can simultaneously serve multiple functions of varying criticality, it cannot easily be labeled as strictly construct relevant or irrelevant. A passage that increases reading load may support the test taker's work in other ways. We also note that identifying context in an assessment task requires more than a surface analysis of its presence or absence. Tasks with a very limited instructional scenario may very well be rich in context, and others, like the task shown in Figure 1, Panel B, may have an instructional scenario that contributes little to the knowledge that the task assesses.

We recognize that both the specific ways that context functions and also their occurrences could vary for a different sets of MKT tasks. This analysis represents only a snapshot of possible context types, the ways in which they are hypothesized to function, and variation of each type. While we have no evidence that particular patterns or lack of patterns would generalize to other measurement situations, we do have evidence from a related study that the patterns are similar when looking at comparable measures in other subjects (Phelps, Howell, & Kirui, 2015).

In conclusion, we think that the approach to describing teaching context in this paper is likely to be useful in better understanding and evaluating MKT task design and as a basis for designing studies that systematically vary the use of context to further explore how those designs function. The analysis illuminates the relation between the types of knowledge a test

taker uses in answering a task, the design of assessment tasks including relevant features of context, and the MKT domain assessed by the task. Explicit attention to the role that context plays in the design of MKT assessments offers the potential to better understand not only the content knowledge that is assessed in particular tasks, but also to begin to develop a theory of how teaching context itself may serve to define this knowledge.

References

- Ball, D. L., & Bass, H. (2002, May). Toward a practice-based theory of mathematical knowledge for teaching. In *Proceedings of the 2002 annual meeting of the Canadian Mathematics Education Study Group* (pp. 3–14).
- Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, 59(5), 389–407. doi: 10.1177/0022487108324554
- Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis*, 25(2), 119–142. doi: 10.3102/01623737025002119
- Gitomer, D., Phelps, G., Weren, B., Howell, H., & Croft, A. (2014). Evidence on the validity of content knowledge for teaching assessments. In T. Kane, K. Kerr, & R. Pianta (Eds.), *Designing teacher evaluation systems: New guidance from the Measures of Effective Teaching project* (pp. 493–528). San Francisco: Jossey-Bass.
- Hawkins, D. (1974). I, thou, and it. In D. Hawkins (Ed.), *The informed vision: Essays on learning and human nature* (pp. 48–62). New York: Agathon Press.
- Herbst, P., & Kosko, K. (2014). Mathematical knowledge for teaching and its specificity to high school geometry instruction. In J. Lo, K. Leatham, & L. Van Zoest (Eds.), *Research trends in mathematics teacher education* (pp. 23–46). New York, NY: Springer.
- Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education*, 39(4), 372–400. Retrieved from <http://www.jstor.org/stable/40539304>
- Hill, H. C., Schilling, S. G., & Ball, D. L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *The Elementary School Journal*, 105(1), 11–30. doi: 10.1086/428763
- Hill, H. C., Sleep, L., Lewis, J., & Ball, D. L. (2007). Assessing teachers' mathematical knowledge: What knowledge matters. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 111–156). Charlotte, NC: Information Age Publishing.
- Howell, H., Phelps, G., Croft, A., Kirui, D., & Gitomer, D. (2013). *Cognitive interviews as a tool for investigating the validity of content knowledge for teaching assessments* (ETS Research Report 13–19). Princeton, NJ: Educational Testing Service.
- Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 17–64). Westport, CT: Praeger.
- Kersting, N. (2008). Using video clips of mathematics classroom instruction as item prompts to measure teachers' knowledge of teaching mathematics. *Educational and Psychological Measurement*, 68(5), 845–861. doi: 10.1177/0013164407313369

- Krauss, S., Baumert, J., & Blum, W. (2008). Secondary mathematics teachers' pedagogical content knowledge and content knowledge: Validation of the COACTIV constructs. *ZDM Mathematics Education*, 40(5), 873–892. doi: 10.1007/s11858-008-0141-9
- McCorry, R., Floden, R., Ferrini-Mundy, J., Reckase, M. D., & Senk, S. L. (2012). Knowledge of algebra for teaching: A framework of knowledge and practices. *Journal for Research in Mathematics Education*, 43(5), 584–615. doi: 10.5951/jresmetheduc.43.5.0584
- McDonald, J. (1992). *Teaching: Making sense of an uncertain craft*. New York: Teachers College Press.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). New York: American Council on Education/Macmillan.
- Mislevy, R., & Haertel, G. (2006). Implications of Evidence-Centered Design for Educational Testing. *Educational Testing: Issues and Practice*, 25(4), 6–20. doi: 10.1111/j.1745-3992.2006.00075.x
- Phelps, G., Howell, H., & Kirui, D. (2015). *Situating the assessment of content knowledge for teaching in the context of teaching practice* (ETS Research Report). Manuscript submitted for publication.
- Phelps, G., Howell, H., Schilling, S. G., & Liu, S. (2015). *Exploring differences in mathematical knowledge for teaching for prospective and practicing teachers*. Manuscript in preparation. *Journal of Mathematics Teacher Education*.
- Phelps, G., Weren, B., Croft, A., & Gitomer, D. (2014). *Developing content knowledge for teaching assessments for the Measures of Effective Teaching Study* (ETS Research Report 14-33). Princeton, NJ: Educational Testing Service. doi: 10.1002/ets2.12031
- Schwab, J. J. (1978). The practical: A language for curriculum. In I. Westbury & N. Wilkof (Eds.), *Science, curriculum, and liberal education: Selected essays* (pp. 287–321). Chicago: University of Chicago Press.
- Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Newbury Park, CA: Sage Publications.
- Tatto, M. T., Schille, J., Senk, S., Ingvarson, L., Peck, R., & Rowley, G. (2008). *Teacher education and development study in mathematics (TEDS-M): Conceptual framework*. Lansing, MI: Teacher Education and Development International Study Center.
- Thompson, P. W. (2015). Researching mathematical meanings for teaching. In English, L., & Kirshner, D. (Eds.), *Third Handbook of International Research in Mathematics Education* (pp. 435–461). London: Taylor and Francis.
- Turner, F. and Rowland, T. (2008) *The knowledge quartet: a means of developing and deepening mathematical knowledge in teaching*. Mathematics Knowledge in Teaching Seminar Series: Developing and Deepening Mathematical Knowledge in Teaching (Seminar 5), Loughborough University, Loughborough.

Interview Prompts to Uncover Mathematical Knowledge for Teaching: Focus on Providing Written Feedback

Yeon Kim (yeonkim10@silla.ac.kr)¹
Silla University, South Korea

Abstract: One area of study that has been gathering enthusiastic attention and interest is mathematical knowledge for teaching (MKT). How to research MKT, however, is still unsettled despite the plethora of unexamined areas of practice. As one of ways to unearth and measure MKT, this study uses interview prompts designed to providing written feedback, as a target area of practice. This study specifies in what ways the interview prompts are used in order to provide a comprehensive method to researching MKT. From interviews across professional communities with different kinds of mathematical expertise, the author develops a conceptual model based on the tasks of teaching, elements of pedagogical context, and domains of MKT. This model provides the fluent character of proficient MKT decision making in teaching practice and explains key features of the design of prompts for investigating and measuring MKT. From the analysis, two claims emerged: bidirectional approaches to investigating MKT and continuous and spontaneous aspects of MKT.

Keywords: mathematical knowledge for teaching, providing written feedback, pedagogical context, measurement, conceptual study.

Introduction

Mathematical knowledge that is specifically connected to the work of teaching has been investigated empirically and theoretically, leading to a significant progression of its conceptualization. In particular, Ball, Thames, and Phelps (2008) developed a practice-based theory of content knowledge for teaching and introduced *mathematical knowledge for teaching* (MKT) as the mathematical knowledge needed to carry out the work of teaching mathematics. Such knowledge has been studied by examining teaching practice, such as job analysis (Ball & Bass, 2003) and by developing its measurement (Hill, Schilling, & Ball, 2004). MKT has been identified as important in teaching mathematics (Lewis & Blunk, 2012) and in student achievement (Hill, Rowan, & Ball, 2005; Rockoff, Jacob, Kane, & Staiger, 2011).

Research on MKT has been conducted with practices of teaching that are prominent in mathematics classrooms, though not all practices. Ma (1999), for example, used four items with some exceptional tasks of teaching and mathematical demands, but did not include all topics and practices in and from teaching. Items developed by the University of Michigan include

¹ This research was supported in part by a grant from the National Science Foundation (Grants DRL-1008317). The author thanks members of the Mathematics Teaching and Learning to Teach Project for their help in developing aspects of this article. Errors are the responsibility of the author.

An earlier version of this research was presented with Mark Hoover at the 2015 KSME International Conference on Mathematics Education and at the 2015 annual meeting of the American Educational Research Association.

substantial tasks of teaching, but not a sufficiently organized approach to mathematical topics and teaching practice.¹ Ball et al. (2008) emphasized a practice-based approach to study content knowledge for teaching. While introducing their conceptualization about such special knowledge, they focused on several and seminal practices of teaching, such as presenting mathematical ideas, responding to students' "why" questions, choosing and developing useable definitions. Nonetheless, the conceptualization has not yet been broadened to explore all practices or all topics in terms of MKT. In other words, these studies do not cover the extensive terrain of mathematical demands in teaching across contexts. For sustainable and elaborated development of the study of MKT, the mathematical demands entailed in teaching now needs to be explored systemically using a clear and comprehensive map of the practice of teaching, mathematics, features of learners, and national or international curriculum or grade levels.

How MKT is studied has yet to receive sufficient attention. Ball and her colleagues studied MKT with a set of analytic tools they developed, using their wide range of experiences and disciplinary backgrounds, for coordinating mathematical and pedagogical perspectives (Ball et al., 2008; Thames, 2009). However, their experiences have not been shared in terms of researching MKT. A robust, reliable, and consistent study of MKT can be expected with an appropriate and good method. Substantial areas of mathematical demands in different practices of teaching still call for investigations from many researchers. Like the collaborative work on the Human Genome Project (International Human Genome Sequencing Consortium, 2001; Naidoo, Pawitan, Soong, Cooper, & Ku, 2011), research on MKT would need cooperative work for elaborate and systematic conceptualization. A major prerequisite for such collective work is the identification of a method to research MKT. If relevant methods of studying MKT are specified, research of MKT will be powerfully advanced. To systemically research MKT, a comprehensive method needs to be specified.

To address the problems, the current study focuses on both using interviews as a comprehensive way to study and measure MKT and researching mathematical demands in providing written feedback, which is an unexamined teaching practice. It does so by building on lessons from the interview prompts used in Ball (1988) and Ma (1999) and on items used to measure MKT (Ball, Bass, & Hill, 2004). Specifically, this paper explores the ways in which interview prompts are developed and used to provide the content and character of MKT, and what, through the use of such interview prompts, might be learned. Particularly, what is entailed mathematically and pedagogically in providing written feedback?

Measuring MKT

Based on Shulman's (1986) notion of pedagogical content knowledge, Ball et al. (2008) specified several subdomains within pedagogical content knowledge (knowledge of content and students [KCS], knowledge of content and teaching [KCT], and knowledge of content and curriculum [KCC]). Further, they identified an important subdomain of "pure" content knowledge unique to the work of teaching, specialized content knowledge [SCK]. SCK is "distinct from the common content knowledge [CCK] needed by teachers and non-teachers alike" (p. 389). They argued that, "teachers' opportunities to learn mathematics for teaching could be better tuned if we could identify those types more clearly" (p. 399).

In earlier research, Ball (1988) created problems that were used in interviews with prospective teachers to explore their proficiency in meeting the mathematical demands of teaching. Ma (1999) extended the use of these interview prompts with practicing teachers in China. Their findings and arguments were critically informed by the carefully designed

mathematical teaching problems represented in the interview prompts. Such mathematical teaching problems are, in short, tools for uncovering mathematics entailed in teaching. Based on analyses of teaching, Ball and Hill's *Learning Mathematics for Teaching* (LMT) project developed multiple-choice items to measure MKT (Ball, Hill, & Bass, 2005; Hill, Ball, & Schilling, 2008; Hill, Schilling, & Ball, 2004). Following the same approach, Herbst and Kosko (2014) developed MKT items for secondary geometry teaching.

To identify key design features of MKT items, my colleagues and I analyzed representations of teaching embedded in MKT assessment items.² We found that the reasoning of item doers was shaped by elements of pedagogical context: student background, teaching purposes, and classroom artifacts, as represented in the teaching situations described in the items. The pedagogical context given in items created situations that required doing mathematics while holding onto a specific pedagogical purpose; the context gave the tasks an MKT, rather than a disciplinary mathematics character. We also found that competent performance on the items depended on reasoning that used, in integral ways, features of the pedagogical context. This conceptualization of pedagogical context offers further ideas about how to design items in ways that measure mathematical knowledge that is fundamentally linked to teaching and not simply disciplinary knowledge that remains remote from effective teaching and learning.

The current study uses open-format prompts designed to investigate the work of teaching that is not readily evident from video of teaching (which was used to design the pedagogical contexts and mathematical tasks of teaching in LMT items). The pedagogical focus of the sub-study reported here is the task of providing written student feedback. This task of teaching requires consideration of the mathematics problem, students' responses to it, and decisions about how best to guide students toward the intended purpose of the problem. It does not mean to suggest that this task of teaching cannot be studied observationally or measured using multiple-choice items. Rather, it needs a means of uncovering more about what was involved in the task and have thus explored the use of designed interview prompts to expand the investigation of MKT. Hence, the current study investigates the following questions: What are the ways in which interview prompts can be developed and used to provide insight into the content and character of MKT? What is entailed, mathematically and pedagogically, in providing written feedback?

Providing Written Feedback

The purpose of feedback is to reduce discrepancies between what is understood and what is intended to be understood (Hattie & Timperley, 2007). Feedback aims to modify a student's thinking to improve his or her learning (Shute, 2008). Apparently, major and complex tasks of teaching involved in the providing of feedback include the sizing up of students' current understandings, directing students to a desired goal, and deciding what information will be provided and in what ways. The complexity of providing feedback is evidenced by a large body of research that encompasses many conflicting findings and no consistent pattern of results. Nevertheless, Hattie and Timperley (2007) reviewed research on feedback and clarified that

... feedback needs to be clear, purposeful, meaningful, and compatible with students' prior knowledge and to provide logical connections. It also needs to prompt active information processing on the part of learners, have low task complexity, relate to specific and clear goals, and provide little threat to the person at the self level. (p. 104)

Shute (2008) made a similar claim that feedback should be nonevaluative, supportive, timely, and specific. Based on a literature review of feedback that concentrates on specific

information to a student about a particular response to a task, Shute (2008) suggested that feedback should address specific features of the student's work with a description of the what, how, and why of a given problem and suggest improvements that the student can manage. This clear feedback would reduce uncertainty in relation to how well students are performing on a problem and what needs to be accomplished to attain the goals.

Feedback, in fact, can promote learning if it is received mindfully (Bangert-Drowns, Kulik, Kulik, & Morgan, 1991). Yeager et al. (2014) emphasized that trust is crucial for successfully delivering written feedback, adding that "mistrust can lead people to view critical feedback as a sign of the evaluator's indifference, antipathy, or bias, leading them to dismiss rather than accept it" (p. 805). Harber et al. (2012) pointed out that many teachers tend to overpraise students for mediocre performance, particularly those subject to negative stereotypes, in order to enhance student self-esteem. However, unlike the teachers' intention, this overpraising in written feedback hinders the development of trust and reinforces minority students' perceptions that they are being viewed stereotypically (Croft & Schmader, 2012; Harber et al., 2012). Feedback should be based on a reflection of a teacher's high standards, not his or her bias, and offer students both an assurance about their potential to reach such high standards and the resources with which they might do so (Yeager et al., 2014). However, written feedback is more unbiased and objective than face-to-face feedback (Kluger & DeNisi, 1996; Shute, 2008).

Critical review about feedback, as specified previously, concentrates on features of feedback that function well to improve students' learning. However, it does not show what tasks of teaching are organically entailed in providing written feedback. The work of teaching includes the activities in which teachers engage and the responsibilities they have to teach content (Ball & Forzani, 2009). The work of teaching occurs in the dynamics initiated by a teacher before, during, and after instruction so as to help students learn the content (Sleep, 2009). Furthermore, in teaching mathematics, specific mathematical goals are critical throughout instruction (Schoenfeld, 2011). Recognizing the varied understandings of the mathematics, for example, probing so as to see what students do or do not know, and responding in ways that address students' errors and that build on student understanding help move students toward goals in instruction. Evidently, providing written feedback is purposeful work: control or modification of a student's ways of thinking or answers to learn something that a teacher desires. Providing written feedback assumes that a teacher gets and synthetically and analytically evaluates signals from students' responses or reasoning, tasks or problems given, and pedagogical situations enacted and created by the teachers and the students. Then, the teacher makes a decision to put forward certain comments to the individual student or to the whole group of students. This description offers a simple glimpse of a feedback loop because it is important to understand the circumstances that result in the differential outcomes and responses of students (Hattie & Timperley, 2007). This loop could be repeated in a lesson until the teacher is satisfied with the control or with the students' revised ideas or answers. Each teacher might have different perspectives and rationales on what feedback works well and what should be highlighted in which situations. However, providing written feedback can be specified with sub-tasks of teaching that are responsive and intrinsic to teaching and entail professional norms of teaching.

Design and Analysis of the Prompts and Interviews

Design of the Interview Prompts

Prompt development started with initial descriptions of the high-leverage practices given in the professionally vetted version on the *TeachingWorks* website.³ The 19 high-leverage practices are tasks central to teaching, which are expected to increase the likelihood for students' learning across a broad range of subject areas, grade levels, and teaching contexts. These high-leverage practices are warranted by research evidence, wisdom of practice, and logic and were developed through many discussions with researchers, expert teachers, faculty members of teacher education, and education policy makers. Of the 19 practices, 16 are crucial in teaching mathematics, as shown in Figure 1. The figure illustrates both apparent interrelationships of the different practices of teaching and gives an overall picture of the practices of teaching mathematics. In other words, Figure 1 represents a comprehensive map of the practice of teaching, which the current study uses to study the mathematical demands entailed in teaching.

Figure 1 shows that providing oral and written feedback is based on assessing students' knowledge. Such an assessment includes selecting and using particular methods to check understanding and monitor student learning and composing, selecting, interpreting, and using information from methods of summative assessment. Providing feedback is influenced by appraising, choosing, and modifying tasks and texts for a specific learning goal. Furthermore, providing feedback may consist of leading a whole-class discussion and setting up and managing small group work. In other words, providing feedback is related to other practices in teaching

Figure 1. High-leverage practices in teaching mathematics

mathematics. Although providing feedback is typical work in teaching, it is crucial to approach this work from a holistic perspective about teaching practice.

Teachingworks explains the providing of oral and written feedback to students on their work as follows:

Effective feedback helps focus students' attention on specific qualities of their work; it highlights areas needing improvement; and delineates ways to improve. Good feedback is specific, not overwhelming in scope, and focused on the academic task, and supports students' perceptions of their own capability. Giving skillful feedback requires the teacher to make strategic choices about the frequency, method, and content of feedback and to communicate in ways that are understandable by students.

This explanation highlights both feedback about a task and processing of a task, which Hattie and Timperley (2007) differentiated. The explanation also emphasizes the need to attend to the student's work and to give specific comments about that work to advance it in a way that they can truly accept and understand (Shute, 2008; Yeager et al., 2014).

To access potentially tacit knowledge and reasoning about what is involved in providing feedback, situations were generated using realistic pedagogical contexts focused on written feedback. On this point, Common Core State Standards were used as a second map to select mathematical topics and determine students' backgrounds. These standards outline a clear set of mathematical skills and knowledge that students will learn in a more organized way both during the school year and across grades. The standards' coherent composition and specific statements offer clear features of mathematical topics and possible information about students' backgrounds that were necessary in creating scenarios for interview prompts. Scenarios of instruction, which would require feedback, and possible student work were sketched. Different elements were considered as well: which elements of pedagogical context are used and how they flow in the scenarios and what mathematical ideas, reasoning, and practices are addressed and how they can be unveiled and interpreted. In deciding each of these elements, the focus was on keeping the situation realistic.

Figure 2 and Figure 3 offer an example that sheds light on the transformation of one initial interview prompt into one used in interviews. In fact, the transformation took place over the course of ten revisions over several meetings with colleagues. The figures illustrate some of the challenges that can occur in developing interview prompts. Both prompts are about providing written feedback as a major practice of teaching, lines of symmetry as a mathematical topic, and fourth-grade students as student background. Nonetheless, Figure 2 does not present a teaching purpose for the instruction; different triangles are used repeatedly without being developed in the instruction; there is no mathematical foundation that interviewees or Mrs. Johnson in the scenario can use to provide feedback; specifically, it is unclear what Mrs. Johnson did pedagogically when drawing symmetry lines on the board; and, at the fourth grade level, the congruency marks offer no indication of whether lines of symmetry exist or not. Figure 3—the final version—specifies first what the teaching purpose is in the provided situation (helping fourth-grade students understand lines of symmetry and how to draw them for two-dimensional figures); it offers a concise and short explanation about the activity that Mrs. Johnson had with her students (using squares and figuring out how to find lines of symmetry); a definition of a line of symmetry is given by the teacher in the scenario, which can be used by the interviewee to provide written feedback in the interview; and student work is provided that includes errors and a

partial understanding about lines of symmetry. Phrases in the scenario have been revised repeatedly to create a succinct and unambiguous scenario that requires written feedback.

Three major questions were decided on: What does the interviewee first notice? What

In a lesson about lines of symmetry for triangles with her fourth grade students, Mrs. Johnson handed out the copies of an equilateral triangle, an isosceles triangle, and a scalene triangle, and students had time to try to fold each triangle into matching parts. Mrs. Johnson then explained lines of symmetry for triangles as lines across the triangles such that the triangles can be folded along the line into matching parts. She also drew symmetry lines on the board as shown in the following:

Then, to assess her students' understanding, Mrs. Johnson gave students time to individually practice drawing lines of symmetry for triangles. When she was monitoring, she noticed that several students drew lines on the textbook as following:

Figure 2. Example of initial interview prompt

Mrs. Johnson wanted to help her fourth-grade students understand lines of symmetry and how to draw them for two-dimensional figures. At the beginning of a lesson, she had the students cut out squares and look for ways to fold a square into two matching parts. The students identified two ways: folding it diagonally and folding it from the middle of one side to the middle of the opposite side. They used rulers to draw lines along their folds. Mrs. Johnson explained that these lines are called lines of symmetry and wrote the following definition on chart paper.

A line of symmetry is a line across a figure such that the figure can be folded along the line into matching parts.

Then, the students did a similar activity with triangles, cutting out triangles and looking for ways to fold each triangle into two matching parts. Here, they noticed that some triangles do not have a line of symmetry.

At the end of this lesson, Mrs. Johnson gave her students a page of practice exercises, where she included three figures that are not triangles and squares. For these figures, one of her students handed in the following:

Figure 3. Example of interview prompt used in actual interviews

written feedback does he/she offer? What would he/she do for a fifteen-minute lesson? These three questions aimed to grasp an interviewee's knowledge and reasoning for providing written feedback. Along with the major questions, several minor questions were added to identify the interviewee's rationale behind his or her responses. Figure 4 includes a list of major and minor questions on the left and on the right one page of possible responses for each question, used to prepare interviews for the interviews. Minor questions helped interviewers prepare for the interviews as well, deciding which minor questions to ask depending on the interviewees' responses. The initial prompts were discussed with colleagues who had teaching experience in grades K-9 and who had studied mathematics and mathematics education. Such discussions were meant to help the researcher anticipate responses and revise the prompts based on whether the prompt would activate and support the task of providing feedback. In other words, collective work was initiated with the designing of interview prompts.

For this sub-study of providing written feedback, interviews were conducted for each interview prompt with two practicing mathematicians, two prospective teachers, and three expert teachers.⁴ Interviewees were carefully selected and contacted. The mathematicians' responses in the interviews offered insight into how disciplinary thinking functions and what appropriately supporting learners means in instruction. The prospective teachers had no prior regular teaching experience and were attending the School of Education working toward their teaching certifications. They were considered a group that needed more professionalism rather than what mathematicians and expert teachers performed in order to investigate mathematical reasoning and knowledge entailed in providing feedback and identify features of MKT in such a task of teaching. To recruit expert teachers, the project team first contacted researchers who were

Grade 4	Geometry [A line of symmetry]	Providing feedback about lines of symmetry
Questions <ul style="list-style-type: none"> • [after getting a response to the first question] How did you figure that out? How do you know? • [after getting a response to the second question] Why did you do that? Why is that? • What did you do to write this feedback? What did you consider? What did you rule out? • [after getting a response to the third question] Why will you plan it? Why is that? <p>[Follow-up questions for Question 2]</p> <ul style="list-style-type: none"> • What does the student appear to know and not to know about lines of symmetry? • Which two-dimensional figures work well with this student's solution to find all lines of symmetry? • What are some other figures where the student might miss some of the lines of symmetry? • What mathematically might be the source of the student's limited understanding? • Did you use or consider this source in your feedback? [If so, ask] How did you use it? [If not, ask] Why didn't you use it? • Do you have any particular mathematical issues or intentions that you considered or emphasized in your feedback? How did you do that? Why are they important? • What do students need to know or what can students do in order to recognize a line of symmetry, identify line-symmetric figures, and draw lines of symmetry? What is important for students to understand about lines of symmetry? • Did you use or consider this knowledge or mathematical importance in your feedback? [If so, ask] How did you use it? [If not, ask] Why didn't you use it? • Would it be helpful for students to be familiar with certain properties of lines of symmetry? • Did you use or consider these properties in your feedback? [If so, ask] How did you use them? [If not, ask] why didn't you use them? • [In the feedback, if the interviewee mentioned the square that the class worked] You pointed out the square in your feedback. Why did you mention it specifically in your feedback? • [In the feedback, if the interviewee mentioned the definition that the teacher wrote down on chart paper] You pointed out the definition in your feedback. Why did you include it in your feedback? There are other possible resources in this classroom. Why didn't you use them? [After hearing the response, and then ask] What do you think about the activity with square? • [If the interviewee did not mention any activities that happened in the provided lesson, ask] How do you think about the activities that given in the lesson? Have you considered them to formulate your feedback? • What would you expect from your feedback? What would this student actually recognize or do after getting your feedback? • Would developing strategies for finding lines of symmetry be the most helpful? • Would it be best to address directly the issue of lines of symmetry passing either through vertices or midpoints of sides? <p>[The following three questions might not needed in the interview, but if the interviewee does not do these things or the interviewer decides to need to know, ask them]</p> <ul style="list-style-type: none"> • Any other more lines of symmetry in the first figure? What about the second and third one? • How do you find them? How do you find a line of symmetry? Are you sure the pairs are matching? • Is there a way to fold a trapezoid into two matching parts? • If a line of symmetry crosses a line of the figure, must it be perpendicular? • How would you show a figure does not have any lines of symmetry? • How would you show that you have all lines of symmetry for a given figure? • What numbers of lines of symmetry are possible? Is there a triangle with just two lines of symmetry? • Are there any types of lines of symmetry that cannot be found in quadrilaterals? <p>[Follow-up questions for Question 3]</p> <ul style="list-style-type: none"> • [If the interviewee did not specify the purpose of this 15-minute lesson, ask] What would be the purpose of this small lesson? • What mathematics would you like to emphasize in that lesson? • [If the interviewee would use some examples of figures, ask] Why did you plan to use these figures? • [If the interviewee does not mention any examples of figures, ask] If you are using examples of figures in this lesson, which would you like to use? Why? 		
<p><i>In reviewing this student's work, what do you notice?</i></p> <ul style="list-style-type: none"> → All the lines drawn are lines of symmetry. They all pass through vertices, not edges. Some lines of symmetry are missing. The missing ones all pass through edges, not vertices. → They seemed to think that all lines of symmetry are lines from a vertex to another vertex. → They seemed to think that lines of symmetry are diagonal. <ul style="list-style-type: none"> • How did you figure that out? How do you know? <p>→ All lines that this student drew pass through vertices, but there is no line between a vertex and a midpoint of a side.</p> <p>[It might be possible to ask additional questions here, but the main investigation in this prompt will be in the second question.]</p> <p><i>Given the lesson just taught, what written feedback might be good to offer this student? Please write your feedback.</i></p> <ul style="list-style-type: none"> → Look at lines of symmetry in your square. One method is folding it diagonally, and another is folding it vertically, or horizontally. Think about lines of symmetry that are from the middle of one side to the middle of the opposite side. → Some lines of symmetry are from the middle of one side to the middle of the opposite side. → All of the lines of symmetry you drew pass through two vertices. If you think about the square example, were there other types of lines of symmetry? Do you see any of those types on these figures? → Why are these lines of symmetry? How do you know if you have them all? <ul style="list-style-type: none"> • Why did you do that? Why is that? [It is possible to ask additional questions here.] <p>→ This student needs to draw lines of symmetry that he/she missed first time. I also hope helping him/her learn how to draw them as the proposed purposes here.</p> <ul style="list-style-type: none"> • What did you do to write this feedback? What did you consider? What did you rule out? [It is possible to ask additional questions here.] <p>→ I used my analysis of the student work. Here, I found this student did not draw a certain type of lines of symmetry. Actually, the class worked with a square in the lesson. So, I used this resource in my feedback.</p> <p><i>If several students had completed the practice exercises as above and there was an opportunity to spend an additional 15-minutes addressing students' misunderstanding, what would you recommend doing?</i></p> <ul style="list-style-type: none"> → The purposes of tomorrow lesson will be same as today's because it is a short lesson and some students did not achieve today's purposes. First, I'll briefly review a page of practice exercises that the students worked today. Here, I'll emphasize all lines of symmetry are not from a vertex to another vertex. Some lines of symmetry are from the middle of one side to the middle of the opposite side, in particular, for an isosceles trapezoid, and a regular hexagon. Second, I'll bring different types of quadrilaterals (rectangles, rhombuses, and isosceles trapezoids) and regular hexagons and octagons, and have the students do similar activities, such as cutting them and finding ways to fold them into two matching parts. It would not take long. Then, I'll bring all figures that we worked on today and tomorrow and arrange them on the board so that the students can apply the working definition that they used today and figure out which figures have lines of symmetry and which figures do not. → For a regular pentagon, a line of symmetry goes from a vertex to the midpoint of a side. → I will bring these figures again on the board and ask some students to draw lines of symmetry and talk about them if they miss some lines. If so, I will ask students to think back the square examples and emphasize lines of symmetry from one midpoint of one side to another midpoint of another side. <ul style="list-style-type: none"> • Why will you plan it? Why is that? <p>→ I would like to help them understand of bisecting an angle or a side as a critical way to draw a line of symmetry.</p> <p>[It might be possible to ask additional questions here.]</p> <p>[Follow-up questions for Question 2]</p> <ul style="list-style-type: none"> • What does the student appear to know and not to know about lines of symmetry? 		

Figure 4. Examples of questions and possible responses

leading research about teaching, teachers, or professional development in the United States. Lists of teachers were gathered based on the researchers' recommendations of expert teachers. They were knowledgeable in mathematics as well as mathematical practice in instructional situations and each had more than ten years of teaching experience. The prospective and expert teachers' responses in the interviews provided insight into what professional reasoning runs in instructional contexts and how it plays out. Observing what these three groups of interviewees did was to help the research team perceive and specify clearly both mathematical and pedagogical reasoning that is engaged in the work of teaching mathematics. Furthermore, interviews were conducted by several interviewers. Each interview was conducted with one interviewee by one interviewer. As previously specified, the current study was collaborative group work, and the collectively performed interviews helped the project team to deepen their insights about mathematical demands entailed in providing written feedback.

Vignette of One Interview and Its Analysis

Each interview was analyzed separately, and then all interviews and analyses were synthetically and analytically probed to investigate and characterize the mathematical demands of providing written feedback. The next section provides a more detailed explanation about the analysis, but this section shows a vignette of one interview and its analysis to illustrate a way of examining and researching MKT.

One interviewee had taught elementary students for approximately 25 years and, for the past 12 years, had worked as a math coach at the K-8th grade level. She also received an M.A. in teaching and learning mathematics. In the interview with the prompts shown in Figure 5, the interviewee immediately recognized that the student in the prompt used a system to make the list. The interviewee noted that the student changed the last two over and over again as well as reordered the first two letters. She was also curious about what made the student decide that the list was complete. The interviewee then wrote down as feedback: "I see you are using a system to find all of the orders. I noticed that your orders are coming in pairs: where the last two groups switch order before the first two groups switch order. How would you know you have all the

<p>To provide an opportunity for his fifth-grade students to use patterns to reason about whether their solution is complete, Mr. Rios assigned the following problem:</p> <p>With four schools participating in the County Honors Chorus, in how many different orders can the school groups be placed on the risers from left to right? The school groups are Talbot, Littlewood, Wyles, and Meadowbrook.</p> <p>Now that students have attempted the problem, he wants to give each student feedback that will support development of a complete list, followed by reasoning that the list is complete.</p> <p>He finds the following student solution:</p> <ol style="list-style-type: none"> 1. TLWM 2. TLMW 3. MWLT 4. MWTL 5. MTLW 6. MTWL 7. TMLW 8. TMWL 9. TWLM 10. TWML 11. LMTW 12. LMWT 13. LTWM 14. LTWM 15. LWTM 16. LWMT 17. WTLM 18. WTML 19. WMLT 20. WMTL 21. WLTM 22. WLMT 	<p><u>Questions by an interviewer</u></p> <ol style="list-style-type: none"> 1. In reviewing this student's work, what do you notice? 2. What feedback might be good to offer this student? Please write your feedback. <ul style="list-style-type: none"> • Why did you decide that? Why is that? [It is possible to ask additional questions here.] • What did you do to write this feedback? What did you consider? What did you rule out? [It is possible to ask additional questions here.] • What does the student appear to know and not to know about making a list systematically? • What might be the source of the student's limited understanding? • Did you use or consider this source in your feedback? [If so, ask] How did you use it? [If not, ask] Why didn't you use it? • Do you have any particular mathematical issues or intentions that you considered or emphasized in your feedback? How did you do that? Why are they important? • What do students need to know or what can students do in order to have a full list of arrangement of elements? • Did you use or consider this knowledge or this ability in your feedback? [If so, ask] How did you use it? [If not, ask] Why didn't you use it? • What would you expect from your feedback? What would this student actually recognize or do after getting your feedback? • [If the interviewee does not do these things, please ask them] I have several mathematics questions. What solutions were missed in the list? • How do you find all orders? How do you find all orders of four different letters? What about five letters? • How would you show that you have all arrangements of the four letters? 3. If you have 15 minutes to close the problem considering this student work, what would you like to do? <ul style="list-style-type: none"> • Why will you plan it? Why is that? • [If the interviewee did not specify the purpose of this 15-minute lesson, ask] What would be the purpose of this small lesson? • What mathematics would you like to emphasize in that lesson?
--	--

Figure 5. One of interview prompt (permutation) developed in the study and questions by an interviewer

choices for arranging the classes?” The interviewee emphasized the importance of the student’s decision point about when he or she finished making a whole list. For an additional 15 minute-lesson, the interviewee planned to start the lesson by asking how many orders there were with W first and then L, for she expected the student to recognize there were six orders each with W and L, but only four orders with M. The interviewee would have asked, “Should there be six with M?” and “What makes you think there should be six?” And the interviewee would then let the student look at what he or she had gotten correct and what he or she had missed. Her reason for planning the lesson in this way was that there was a mistake in the middle of the list, but, at the bottom of the list, orders with W and L were listed very systemically. Because the student seemed to have a better sense of making a list with L and W, the interviewee expected that the student could figure out the regularity in orders with W and L and apply it to have a complete list.

The interviewee’s responses were analyzed in terms of the mathematical work of teaching engaged in providing written feedback. First, there were shown several sub-tasks of teaching related to providing written feedback. (i) Identifying a mathematical feature that needs to be emphasized in the situation—the interviewee immediately recognized the main mathematical topic of this situation was having a complete list to show all the orders of the four schools. Throughout the interview, although the interviewee used different terms, such as “organization,” “regularity,” “system,” and “efficient,” she emphasized a systematic way to make a complete list with any number of cases from patterns involved in this situation. (ii) Playing with ways in which the student’s work is produced from the given problem—the interviewee investigated the student’s work and recognized that there is a systemic way to make the list. The interviewee also acknowledged that this student had some sense about permutation, but not enough to make a whole list. The interviewee also reflected on what this student’s orders would be if there were two or three schools because the student might be able to make a generalizable list from the smaller example. (iii) Formulating critical feedback—rather than correcting the student’s work or directly specifying what to do next, the interviewee wanted to give something that made the student think more broadly and deeply. She also clarified two elements in her feedback. One was what the teacher noticed in the student’s work and the other was a question to help the student reflect on his or her work and reasoning and investigate and find a way to build a complete list.

The mathematical knowledge and reasoning entailed in providing written feedback were also analyzed. The interviewee was very confident that the most important thing in this situation was letting the student recognize whether or not he or she had all the orders and helping the student understand a generalizable way to make a full list from patterns involved in this situation. Her responses and comments were consistently geared toward the student. However, the interviewee did not ascertain what orders the student had missed until the interviewer asked about it. At the beginning of the interview, she briefly recognized that the orders starting with L and W were systematically listed, but those starting with T and M were not. Rather than just fixing the student’s work, the interviewee put more value on mathematical generalization and working with the student’s reasoning toward it. It was also interesting that the interviewee recognized completeness as a critical issue in this situation. However, the interviewee did not seem to know the total number of permutations using combinatorics ($4!$), but she could explain her reasoning to get a full list of orders by using a tree diagram.

The final thing examined in the analysis of this interview was distinctive mathematical knowledge and reasoning for teaching. The interviewee recognized that the last entries in the

student's list were mathematically well organized, but she did not use it directly in her feedback. Rather, she focused on reasoning related to mathematical completeness and mathematical generalization to approach this situation, and she seemed to believe that mathematics instruction needed to concentrate on them rather than on merely getting correct answers. This might be the reason why she did not rush to solve the provided problem. She investigated the student's work and tried to specify possible reasoning that the student might have used in the work. Her feedback was specific enough to make the student think about what he or she had done. In general, her sense about providing written feedback is to give the student a chance to reflect on his or her work based on what the teacher noticed in the work rather than correcting the work.

Analysis of the Interviews

Each interview was summarized and analyzed carefully by the interviewer using the following questions: How did the interviewee respond to each question of the prompt? What was the rationale for each response? What is the mathematical work of teaching as it relates to this teaching situation as suggested through this interview? What mathematical knowledge and reasoning are entailed in this work as suggested through this interview? What does this interview suggest about what might be distinctive about mathematical knowledge and reasoning for teaching? During this time, we continually revised the interview prompt so as to create with clear language realistic situations.

The summary notes were major resources for analyzing the data. During research meetings, the interviewers shared their experiences in the interviews and discussed major interests of the current research using the questions introduced in the previous paragraph. Figure 6 and Figure 7 show two summary notes, one from an interview conducted with an expert teacher and one from an interview conducted with a mathematician. One of the predominant features shown in the interview with the expert teacher was the teacher figuring out and using information from the pedagogical context. The mathematician, on the other hand, merely checked the correctness of the answer, neglecting to dig in and create feedback in the provided situation. In other words, the student work was mathematically analyzed by the expert teacher using pedagogical considerations of the teaching purpose; her analysis was greatly used to formulate feedback. In that situation, the bottom orders in the student's list constitute a critical clue that formulates feedback according to the given teaching purpose, "use patterns to reason about whether their solution is complete." Checking the correctness of the answer is insufficient to formulate feedback. This task of teaching requires complicated sub-tasks of teaching that are mathematically and pedagogically delicate and that also require comprehensive analyzing and decision making.

This research, again, does not aim to characterize the three different groups of interviewees in the context of providing written feedback. However, interviews with prospective teachers presented challenges to approaching the given situation and making a decision. Their reasoning, as revealed in the interviews, was inconsistent. Some prospective teachers only dug into the problem used in the given situation, just like learners would. Some prospective teachers solved the problem while others did not. Most of them recognized some information about the pedagogical context. Some of them used it to formulate feedback while others did not. It seems clear that one thing that is critical in providing feedback is seizing and analyzing mathematical and pedagogical information simultaneously, immediately and synthetically. It is a real challenge, however, to gather such information. Furthermore, interviewees in each group did not always respond to the prompts in the same way. In other words, mathematical, pedagogical, partial,

analytical, and synthetic tendencies and professionalism for figuring out the provided situation and making a decision were different across participants.

MKT Interview Summary Field Notes		
Date of Interview: [REDACTED] Researcher: [REDACTED] Interviewee Type: Expert Teacher Files: [REDACTED]		
Description of Response and Rationale		
Summarize the participant's responses to the prompt and its questions in the table below. Use the following questions to guide your responses 1. How did the interviewee respond to each question/request of the prompt? If short, record the response; otherwise summarize the response in a sentence or two. 2. What was the rationale for each response?		
Question/Request	Response	Rationale
In reviewing this student's work, what do you notice?	The participant immediately recognized that there is an organization to list letters. She specified that the student changed the last two over and over again as well as reordered the first two letters. She was also curious about what makes the student decide that the list is completed.	This student has a system to make a list, and the teacher needs to understand what the system is, and what regularity the student used.
What feedback might be good to offer this student? Please write your feedback.	I see you are using a system to find all of the orders. I noticed that your orders are coming in pairs: where the last two groups switch order before the first two groups switch order. How would you know you have all the choices for arranging the classes?	The participant identified that the feedback consists of two parts: what the teacher notices in the student's work and getting information about the whole process of solving the problem. What she first does is says what she sees in the student's work. If her observation is different from what the student intended, the student will let her know his/her original intention, or he/she could recognize what the teacher or the other sees in his/her work. Furthermore, in this situation, it is important to know the student's decision point about when he/she
Reflection on the Mathematical Work of Teaching		
Reflect on the interview and consider what you may be learning in relation to the work of teaching in the teaching situation and its mathematical demands. Please record your thoughts in relation to the following three questions. 3. What is the mathematical work of teaching related to this teaching situation as suggested by this interview? What did this person do in responding (conceived as work		
<p>of teaching)? What did the person's work suggest about the mathematical work of teaching? (one paragraph)</p> <p>(i) Identifying a mathematical feature that needs to be emphasized in the situation: The participant immediately recognized the main mathematical topic of this situation is having a complete list to show all orders of four schools. Throughout the interview, although the participant used different terms, such as organization, regularity, system, and efficient, she emphasized a systematic way to make a complete list with any number of cases from patterns involved in this situation. (ii) Playing with ways in which the student's work is produced from the given problem: the participant investigated the student's work and recognized that there is a systemic way the student used to make a list. The participant also acknowledged that this student had some sense about permutation, but not enough to make a whole list. The participant also reflected that if there are two or three schools, what this student's orders would be because the student might be able to make a generalizable list. However, it does not seem that this participant critically use the possible sources of the student's misunderstanding. (iii) Formulating critical feedback: rather than correcting the student's work or directly specifying what to do next, the participant would like to give something that makes the student think more broadly and deeply. She also clarified the two elements as feedback. One is what the teacher noticed in the student's work and the other is a question to help the student reflect on his/her work and reasoning and investigate and find a way to build a complete list.</p> <p>4. What mathematical knowledge and reasoning are entailed in this work as suggested by this interview? (one paragraph)</p> <p>The participant was very confident that the most important thing in this situation is letting the student recognize whether or not he/she had all orders and helping the student understand a generalizable way to make a full list from patterns involved in this situation. Her responses and comments were consistently geared toward them. However, the participant did not investigate exactly what orders the student missed until before the interviewer asked it. At the beginning of the interview, she briefly recognized that the orders starting with L and W are systematically listed, but not those starting with T and M. Rather than just fixing the student's work, the participant put more value on mathematical generalization and working with the student's reasoning toward it. It was also interesting that the participant recognized completeness as a critical issue in this situation. However, the participant did not seem to know the total number of combinations based on combinatorics (4!) very well, but she could explain her reasoning to get a full list of orders by using a tree diagram.</p> <p>5. What does this interview suggest about what might be distinctive about mathematical knowledge and reasoning for teaching? (This might be about something the interviewee evidenced, or about something noticeably absent or different from what would be expected.) (one paragraph)</p> <p>The participant exactly recognized that the last entries in the student's list are mathematically well organized, and she did not use it directly in her feedback. Rather, she focused on reasoning related to mathematical completeness and mathematical</p>		

Figure 6. Example of summary notes of one expert teacher

MKT Interview Summary Field Notes		
Date of Interview: [REDACTED] Researcher: [REDACTED] Interviewee Type: Mathematician Files: [REDACTED]		
Description of Response and Rationale		
Summarize the participant's responses to the prompt and its questions in the table below. Use the following questions to guide your responses 1. How did the interviewee respond to each question/request of the prompt? If short, record the response; otherwise summarize the response in a sentence or two. 2. What was the rationale for each response?		
Question/Request	Response	Rationale
1. In reviewing this student's work, what do you notice?	I would take all the Ts. And then, the others are systematically ordered. Some organized ways are tried to make orders. I like an organized way to approach the problem. ... There are Ts until 10 th . I can see all Ts.	I don't know how to handle it. ... Understanding what the question is always the hardest part. If there is an error, where does it come from? (And then, responded to this question)
2. What feedback might be good to offer this student? Please write your feedback.	Go back and check this out. ... Look systematically to work this out. ... (After responding for the question about what this student knows, the interviewer ask again this question) Try to do an organizing way. All the possibility to do with Ts at the first. That's all you have to do. Are there anything that does not fit? Are there anything that you missed?	I don't know what to do. I don't want to interrupt. This should be 24, not 22. I would not ask a formula. I don't want to say binomial. I think this student know organizing all the Ws in the first column, but not the other ones. Ls are pretty good, but not the Ts. Because of the lack of confidence, this student missed the orders. ... Let students play with the problem rather than teaching how to solve it. I expect that the student is confident to solve the problem quickly.
3. If you have 15 minutes to close the problem considering this student work, what would	I don't know. I really do not know. ... I might go through with students. But, let students	This is a so detailed question. ... I don't want to say if you do something, it
5. What does this interview suggest about what might be distinctive about mathematical knowledge and reasoning for teaching? (This might be about something the interviewee evidenced, or about something noticeably absent or different from what would be expected.) (one paragraph) Catching up the teaching purpose in the provided situation, analyzing the student work, noticing the student error and using student work to formulate feedback may be a distinctive feature of mathematical knowledge and reasoning for teaching. Furthermore, this knowledge may contain that the idea about a systemic approach is part of the content that needs to be taught to students. When asked about giving another problem, he agreed using another problems, but did not come up with appropriate problems (easy, difficult, complex, etc.). This seems to suggest that being aware of appropriate additional problems is distinctive about knowledge for teaching.		

you like to do?	do this problem and argue each other. That would be wonderful. ... I might have some conversation with students, but I don't know.	would be come up. ... I don't know because I have never been there. I don't have any strong opinion. ... I would do with similar problems. I mean examples are wonderful. (The interview asks "Do you have any particular example?") No. I don't have any now. (The interviewer asks "Would you use complex or easy one?") I don't know.
Reflection on the Mathematical Work of Teaching		
Reflect on the interview and consider what you may be learning in relation to the work of teaching in the teaching situation and its mathematical demands. Please record your thoughts in relation to the following three questions.		
3. What is the mathematical work of teaching related to this teaching situation as suggested by this interview? What did this person do in responding (conceived as work of teaching)? What did the person's work suggest about the mathematical work of teaching? (one paragraph) The interviewee seems to immediately recognize 22 orders is a wrong answer. However, analyzing student work or creating specific feedback were not shown. He strongly believed the power of students' argument rather than teaching particularly how to solve the problem, but did not provide any guidance to help the student move forward in his/her thinking. The teacher must recognize the error that the student is making. They must next be able to attribute the error to the student's thinking or reasoning. They must then be able to provide feedback that addresses the error in thinking or reasoning in a way that is actionable by the student. Several of these pieces were missing in this participant's response. For example, in his feedback, this participant asked the student to go back and check again, but did not provide any guidance to help the student move forward in their thinking.		
4. What mathematical knowledge and reasoning are entailed in this work as suggested by this interview? (one paragraph) The interviewee thought that 22 orders was incorrect and the orders were not well organized in the top ones. However, he did not try to figure out which orders the student was missing. Part of the mathematical work of teaching seems to be finding the correct solution. It seems like the mathematical knowledge and reasoning involved in this work includes being able to determine the understanding underlying a mistake and connect it to the student's broader thinking and reasoning.		

Figure 7. Example of summary notes of one mathematician

Based on investigating the distinctive features of mathematical knowledge and reasoning for teaching as it relates to providing written feedback using the interviews with the three different groups, my colleagues and I were able to build up our understanding of the main features of MKT entailed in providing written feedback. Throughout the analysis and discussion

of summary notes, features related to confident and professional reasoning and performance in providing feedback were discovered. The first feature concerned whether or not interviewees had similar teaching experiences to those presented in the situation provided, showing that coherent and logical reasoning exists for approaching and examining provided instructional situations and making a decision based on the teaching purposes embedded in the situation. Confidence about such reasoning was critical as well. The teacher's confidence seemed to determine how well he or she explained the rationales of the analysis and decision making.

The second feature was the explicit recognition that providing feedback aimed to help extend students' understanding by offering specific advice rather than correcting their responses. The appropriateness of specific comments was also considered because comments that were too particular could reduce the chances that students would improve their response. Comments that were too vague offered no help to students.

The third feature was recognizing the instructional purpose in the provided situation and sticking to that purpose while investigating information, formulating feedback and specifying the rationales of such feedback. For example, the teaching purpose in the prompt shown in Figure 5 is to provide "an opportunity for his fifth-grade students to use patterns to reason about whether their solution is complete." This purpose includes pedagogical concerns (e.g., providing an opportunity for fifth-grade students) and mathematical topics (e.g., patterns and mathematical completeness). Immediate recognition of the purpose is prominent at the beginning of approaching the provided situation, and this recognition works critically throughout the providing of feedback.

The fourth feature is analytically and synthetically recognizing information from the pedagogical context. Figuring out and mathematically analyzing student work was critical. This task includes probing the reasoning behind the student work and having reasonable assumptions to explain what gives rise to students' wrong responses. Furthermore, other tasks related to recognizing information included looking for instructional resources, evaluating them to formulate feedback, and deciding which one should be used to create feedback. Identifying mathematical content areas related to the provided mathematical topic was important, too.

Based on the features discovered in the providing of feedback on a task of teaching, my colleagues and I have tried to develop a consistent and logical framework that illustrates proficiency in providing feedback in terms of sub-tasks of teaching. In the process of developing the framework, different domains of MKT used in the particular task emerged. Sub-domains of MKT are categorized by different features of "mathematical knowledge needed to perform the recurrent tasks of teaching mathematics to students" (Ball et al., 2008, p. 399). For example, CCK is the mathematical knowledge and skill used in settings other than teaching, such as "simply calculating an answer or, more generally, correctly solving mathematics problems" (Ball et al., 2008, p. 399). CCK seemed dominant in the mathematician's reasoning in the situation of formulating feedback while SCK, KCT and KCS seemed prevalent in the expert teachers' reasoning. More interestingly, while SCK seemed to be extensively implemented to investigate the mathematics embedded in the provided situation, KCT and KCS seemed to be at work when figuring out the pedagogical context and making a decision regarding the student in the provided situation. CCK seemed to function among SCK and KCS, but primarily garnered attention in identifying that the sub-domains of MKT overlap, through consecutive sub-tasks of teaching.

It was also discovered that different aspects of pedagogical context function across different tasks of teaching. For example, identifying the instructional purpose was closely related to teaching purpose and mathematical topic, and investigating student work was connected to mathematical topic. Student background, such as fifth-grade students, was significant to formulate feedback.

The mathematical work associated with the task of providing written feedback was iteratively analyzed using the interviews as empirical grounding, as explained in this section. The current study used an interactional conceptualization of teaching, as described in Cohen, Raudenbush, and Ball (2003) and a practice-based conceptualization of MKT by Ball and Bass (2003). In other words, rather than trying to describe the MKT held by interviewees or the particular ways in which they reasoned about the prompts, I was trying to use the ways in which they reasoned about the prompts to characterize professionally defensible knowledge and practice. The research questions for the larger study are as follows.

1. What is the mathematical work of teaching in unexamined areas of practice?
 - a. areas perhaps not readily studied using video
 - b. specific practices not previously studied
 - c. key areas with distinctive MKT demands (such as impromptu talk)
2. What mathematical knowledge and reasoning are entailed in this work?
3. What is distinctive about mathematical knowledge and reasoning for teaching?
4. What are key features of the design of prompts and interview methods for uncovering MKT?

The methods of analysis for the study reported here are consistent with the job analysis and conceptualization described by Thames (2009), but are applied to the interview data rather than to video.

Revision of the Prompts

Based on the analysis of the interview, the prompts were both revised and analyzed. The first major question—what does an interviewee notice first—was written in the document. It was decided, however, that the question should be removed and asked orally. The printed question seemed to force the interviewee to focus on noticing things within the provided situation rather than allowing the interviewee to get started in the given situation and find whatever the interviewee happened to pick up on. Moreover, conducting interviews helps the project team examine the prompts, specifically how providing written feedback is decomposed, what mathematical demands are entailed in each task of teaching, and which elements of pedagogical contexts function in each task of teaching. Although the prompts to investigate MKT were developed by the author and the project team, conducting and analyzing the interviews with the developed prompts helped extend our understanding of MKT as well. The characteristics found by the iterative process are specified in the following section.

What is Entailed in Providing Written Feedback

From the iterative and synthetic analysis of the interviews and the interview prompts, it was found that skillfully providing students with written feedback requires teachers to draw on purpose and relevant information given in the pedagogical context and flexibly use knowledge resources across different domains of MKT. MKT seems to involve an ability to recognize important and adequate information about pedagogical context in teaching to make a decision about which actions to perform. To characterize this distinctive knowledge and reasoning with

competent performance of the task, this study uses three features: tasks of teaching, pedagogical context, and domains of MKT.

Sub-Tasks of Teaching

To develop a specification of the mathematical demands of providing feedback, the sub-tasks of teaching entailed in responding to the prompt were analyzed first. Specifying the work of teaching mathematics is critical to both characterize teaching in terms of the dual foci of mathematics and instruction that Ball (1993) specified as the nature of teaching mathematics. Providing written feedback is not just one simple task, but includes several combined tasks integral to developing feedback. Characterizing the work of teaching is a critical step in identifying MKT, as is specifying what this work is and what it requires. I found competently providing feedback, expressed in general terms, involves four sub-tasks:

1. tracking on the instructional purpose of the problem and/or when in the students' learning trajectory the problem is being used, and what that implies for the mathematical territory of the problem;
2. making sense of the student work in relation to the instructional goals, the mathematical structure and territory of the problem, and the multiple ways that the problem can be approached;
3. identifying resources in the student work in relation to helping the student recognize the need for further work and have a way to make further progress;
4. deciding on a clear aim for the feedback and using the resources to design feedback that supports students in being able to work toward the instructional goal.

These sub-tasks tend to unfold in a linear fashion, but each sub-task includes a non-linear fashion of tasks. Furthermore, the sub-tasks may unfold flexibly and may cycle. These can also be expressed in more particular terms for the given scenario in each specific prompt.⁵ This general description of providing written feedback characterizes that the tasks of teaching include the providing of written feedback.

Role of Pedagogical Context

As previously identified, one component that is salient to reasoning in teaching practice is pedagogical context. My analysis suggests four elements that can support MKT reasoning about providing feedback. These elements consist of the instructional purpose given in the scenario, the mathematical topic discussed in the provided instruction, instructional resources provided to support instruction, which can be used in instruction, and student background, which can offer information for feedback. These elements of pedagogical context provide support for reasoning about providing feedback. They might also be suited for other tasks of teaching.

Another issue is that these four elements of pedagogical context do not operate in all four sub-tasks simultaneously. In other words, there are targeted or untargeted elements in each aspect of the work. In the first sub-task, teaching purpose and mathematical topic are used to identify the instructional purpose and to consider the mathematical structure. Second, the mathematical topic and instructional resources are used to draw a map between the student work and the original problem and resources used in instruction. Teaching purpose is briefly used to interpret how well the student work matches the teaching purpose. Third, mathematical topic and instructional resources are reconsidered, but their uses are different. Here, they are used so as to enable the student to recognize the need for further work. All elements are used in different

moments in the different sub-tasks, which are elaborated in the mathematical work of teaching, to formulate feedback. In particular, the instructional purpose provides direction for feedback.

Overlapping Sub-Domains of MKT

In parallel to the analysis of elements of pedagogical context, I examined each sub-task in terms of the domains of MKT. Each sub-task involves different demands in relation to MKT's sub-domains. First, to recognize the mathematical structure and identify resources, SCK and KCT play a role in identifying the instructional purpose, establishing core ideas of the provided problem, and determining which resources were mathematically used. The second sub-task entails the use of SCK, CCK, and KCS to identify relationships between the problem focused on and student work, solve the problem, identify which resources were used and how they were used in student responses, and generate possible reasons for the student work. The third sub-task requires SCK to find resources used to suggest ways to make further progress. The fourth sub-task involves SCK, KCS and KCT to encourage the student to review and develop her or his work. Providing written feedback entails SCK, CCK, KCS and KCT, though the use of these sub-domains shifts across its subtasks. SCK plays a major role, but CCK, KCS and KCT are also critical in providing feedback.

Figure 8 synthetically provides a view of the distinctive character of competent performance of the task conceptualized in relation to three basic features: the tasks of teaching, the pedagogical context, and the domains of MKT. This diagram, which is chronological from top to bottom, has a major axis with the sub-tasks of teaching for providing written feedback. Each task also entails different elements of the pedagogical context, which are represented by

Figure 8. Proficiency in providing written feedback

dark or light gray rectangles depending on the targeted or untargeted elements. Furthermore, this diagram shows which different domains of MKT functions in each task of teaching, which are represented by abbreviations in the arcs.

Discussion

This study aims to investigate the ways in which interview prompts are developed and used to provide the character of MKT. It also explores what is entailed mathematically and pedagogically in providing written feedback to demonstrate this type of method for studying MKT. From the analysis of the interview prompts and the interviews conducted with people who have different experiences related to teaching mathematics, the current study sets up the three interrelated aspects of mathematical demands: general description of sub-tasks of teaching entailed in providing written feedback, multiple and selective use of elements of pedagogical context in each sub-task, and continuous and simultaneous functions of different domains of MKT. Furthermore, this study specified how the interview prompts were used as a comprehensive way of studying MKT. Based on the analysis above, two issues have emerged related to investigating MKT and aspects of MKT.

In both generating and modifying prompts and probing interviews, prominence is given to the bidirectional approaches to investigating MKT. Developing prompts first aimed at creating situations that provided space for interviewees to organically provide written feedback. The development made certain assumptions about what providing written feedback involves, how interviewees might respond to prepared questions, and what reasoning would prevail. Because interview prompts touch on specific mathematical areas, the mathematical features embedded in the prompts were also carefully considered, including the mathematical facts, practices, and reasoning that would be related to teaching and learning in these particular instructional situations. Developing interview prompts entails analyzing the prompts themselves with such questions. Hypothesized ways of addressing the developed situations mathematically and pedagogically were set up and used to trim and elaborate the prompts. The prompts were designed for the interviews, but the process of creating the prompts also shed light on the instructional situations embedded in the prompt.

Conducting interviews aims to examine reasoning used in a given situation. Interviews with the three groups of people offered a sense of the pedagogical and mathematical reasoning at play in the given instructional situation. Although possible responses were prepared when developing the interview prompts, the interviews provided the project team with both a finer sense of how to revise the interview prompts and a better awareness of aspects of MKT entailed in providing written feedback than merely the step of developing the prompts. Conducting and analyzing interviews enabled the project team to scrutinize the interview prompts and trim them extensively. Going back and forth between creating and revising interview prompts and conducting and probing interviews ultimately unveiled the distinctive characteristics of MKT. This bidirectionality is also used as a way to design items and examine teaching (Jacobson, Remillard, Hoover, & Aaron, in press), and a common feature of MKT research. An interview is an efficient way of targeting and investigating a particular work of teaching in studying MKT. Creating the situation for an interview, however, requires intentional effort unlike the use of video clips, which show instruction clearly without requiring effort to create certain instructional situations. In this sense, the use of interviews in this study is not typical to MKT research and requires close attention to the bidirectionality between developing interview prompts and analyzing interviews.

The second major claim made by this study is that the features of MKT are continuous and simultaneous rather than separate and isolated. Ball (1993) described the nature of pedagogical deliberations and claimed to have understood more about the processes of pedagogical and mathematical deliberation in teaching mathematics. Shulman (1987) also claimed to have followed the process of pedagogical reasoning and actions that is used in teaching. He emphasized the continuity of the process of reasoning with the identification that “pedagogical reasoning and action involve a cycle through the activities of comprehension, transformation, instruction, evaluation, and reflection” (p. 14). These researchers pointed out the process in reasoning and performance engaged in teaching. The current study claims not only continuity in each of three different features but also simultaneousness among them.

To specify the complicated nature of teaching as an elaborated process in terms of tasks of teaching, domains of MKT, and pedagogical context, the current study suggests the three features as a way of describing the deliberation for professional and competent performance of providing written feedback, and to conceptually elaborate MKT. However, this does not mean that the three features operate independently. The three features are closely related to moment-to-moment teaching practice because the deliberations in teaching practice are continuous. Furthermore, elements of each feature functions continuously rather than separately or absently. Sub-tasks of teaching are continuously performed and elements of pedagogical contexts are always functioning. Different domains of MKT functions continuously and simultaneously. Different elements of the pedagogical context are continuously drawn on. This means that each feature includes continuity as well.

Teaching is intricate (Ball & Forzani, 2009). Therefore, investigating MKT requires a careful, analytic, and clear method to deepen and widen the study of nature of MKT by researchers interested in it. Also, an insightful lens is needed to scrutinize its intricacies and clarify the distinctive features of teaching mathematics in terms of MKT. Bidirectionality between the interview prompts and the interview as an approach to researching MKT offers a microscope by which we can discern the critical characters of MKT and teaching. As analysis using the bidirectional approach, the current study also claims that continuity and simultaneousness are major features among the three features used in this study to conceptualize the competent performance of providing written feedback. However, one of major limitations of the current study is that the method for studying MKT illustrated here requires use and validation by multiple groups of researchers to get universality as a robust method for studying MKT.

Conclusions

This study has focused on providing written feedback and investigating key features in the design of prompts for uncovering MKT. This study argued that three dimensions—the decomposition of the task of providing feedback, elements of the pedagogical context, and sub-domains of MKT—are useful in characterizing the distinctive MKT reasoning involved in providing feedback and in designing pedagogical scenarios that can be used as part of a wide variety of tools for engaging, studying, and measuring MKT. This study has found the conceptual tools described here helpful across other datasets but there is a need to continue the analysis across other tasks of teaching and other data sources to both test and refine these initial ideas.

References

- Ball, D. L. (1988). *Knowledge and reasoning in mathematical pedagogy: Examining what prospective teachers bring to teacher education*. Unpublished doctoral dissertation, Michigan State University, Lansing.
- Ball, D. L. (1993). Halves, pieces, and twoths: Constructing representational contexts in teaching fractions. In T. P. Carpenter, E. Fennema & T. A. Romberg (Eds.), *Rational numbers: An integration of research* (pp. 157–196). Hillsdale, NJ: Lawrence Erlbaum.
- Ball, D. L., & Bass, H. (2003). Toward a practice-based theory of mathematical knowledge for teaching. In B. Davis & E. Simmt (Eds.), *Proceedings of the 2002 annual meeting of the Canadian Mathematics Association Study Group* (pp. 3-14). Edmonton, AB: CMESG/GDEDM.
- Ball, D. L., Bass, H., & Hill, H. C. (2004). Knowing and using mathematical knowledge in teaching: Learning what matters. In A. Buffler & R. Laugksch (Eds.), *Proceedings of the twelfth annual conference of the South African Association for Research in Mathematics, Science and Technology Education (SAARMSTE)* (pp. 51–65). Durban, South Africa: SAARMSTE.
- Ball, D. L., & Forzani, F. (2009). The work of teaching and the challenge for teacher education. *Journal of Teacher Education*, 60(5), 497–511.
- Ball, D. L., Hill, H. C., & Bass, H. (2005). Knowing mathematics for teaching: Who knows mathematics well enough to teach third grade, and how can we decide. *American Educator*, 29(3), 14–17, 20–22, 43–36.
- Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, 59(5), 389–407.
- Bangert-Drowns, R. L., Kulik, C. C., Kulik, J. A., & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, 61(2), 213–238.
- Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis*, 25(2), 119–142.
- Croft, A., & Schmader, T. (2012). The feedback withholding bias: Minority students do not receive critical feedback from evaluators concerned about appearing racist. *Journal of Experimental Social Psychology*, 48(5), 1139–1144.
- Harber, K. D., Gorman, J. L., Frank, P., Butsingh, S., Tsang, W., & Ouellette, R. (2012). Students' race and teachers' social support affect the positive feedback bias in public schools. *Journal of Educational Psychology*, 104(4), 1149–1161.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Herbst, P., & Kosko, K. (2014). Mathematical knowledge for teaching and its specificity to high school geometry instruction. In J. J. Lo, K. R. Leatham & L. R. Van Zoest (Eds.), *Research trends in mathematics teacher education* (pp. 23–45). Switzerland: Springer.
- Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education*, 39(4), 372–400.

- Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, 42(2), 371–406.
- Hill, H. C., Schilling, S. G., & Ball, D. L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *The Elementary School Journal*, 105(1), 11–30.
- International Human Genome Sequencing Consortium. (2001). Initial sequencing and analysis of the human genome. *Nature*, 409(6822), 860–921.
- Jacobson, E., Remillard, J., Hoover, M., & Aaron, W. (in press). The interaction between measure design and construct development: Building validity arguments. In A. Izsák, J. T. Remillard & J. Templin (Eds.), *Psychometric methods in mathematics education: Opportunities, challenges, and interdisciplinary collaborations. Journal for Research in Mathematics Education monograph series*. Reston, VA: National Council of Teachers of Mathematics.
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254–284.
- Lewis, J., & Blunk, M. (2012). Reading between the lines: Teaching linear algebra. *Journal of Curriculum Studies*, 44(4), 515–536.
- Ma, L. (1999). *Knowing and teaching elementary mathematics: teachers' understanding of fundamental mathematics in China and the United States*. Mahawah, NJ: Lawrence Erlbaum Associates.
- Naidoo, N., Pawitan, Y., Soong, R., Cooper, D., & Ku, C. (2011). Human genetics and genomics a decade after the release of the draft sequence of the human genome. *Human Genomics*, 5(6), 577–622.
- Rockoff, J. E., Jacob, B. A., Kane, T. J., & Staiger, D. O. (2011). Can you recognize an effective teacher when you recruit one? *Education Finance and Policy*, 6(1), 43–74.
- Schoenfeld, A. H. (2011). Toward professional development for teachers grounded in a theory of decision making. *ZDM*, 43(4), 457–469.
- Shulman, L. S. (1986). Those who can understand: Knowledge growth in teaching. *Educational Research*, 15(2), 4–14.
- Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1–22.
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153–189.
- Sleep, L. (2009). *Teaching to the mathematical point: Knowing and using mathematics in teaching*. Unpublished doctoral dissertation, University of Michigan, Ann Arbor.
- Thames, M. H. (2009). *Coordinating mathematical and pedagogical perspectives in practice-based and discipline-grounded approaches to studying mathematical knowledge for teaching (K-8)*. Unpublished doctoral dissertation, University of Michigan, Ann Arbor.

Yeager, D., Purdie-Vaughns, V., Garcia, J., Apfel, N., Brzustoski, P., Master, A., . . . Cohen, G. L. (2014). Breaking the cycle of mistrust: Wise interventions to provide critical feedback across the racial divide. *Journal of Experimental Psychology: General*, 143(2), 804–824.

Footnotes

¹ Their items mainly concentrated on the following practices of teaching in algebra and geometry for elementary levels: evaluating understanding, choosing examples – illustrating a concept, evaluating explanations, choosing representations, evaluating difficulty, choosing examples – selecting a problem for an exercise, and figuring out non-standard work.

² I would like to acknowledge the contribution of the MTLT project, particularly, Yvonne Lai, Erik Jacobson, and Mark Hoover for outlining the foundation of this work.

³ Here is the link: <http://teachingworks.soe.umich.edu/work-of-teaching/high-leverage-practices>

⁴ I would like to thank the MTLT project, particularly, Rachel Snider, Lindsey Mann, Joy Johnson, and Mark Hoover, for conducting interviews.

⁵ Appendix shows particular description engaged in the interview prompts shown in Figure 5.

Appendix

Particular description of the work of providing written feedback in the interview prompt (permutation), shown in Figure 5

1. Tracking on the purpose of engaging students in using patterns to reason about the completeness of a solution to a permutation problem.
2. Working through the student solution, imagining how the student was likely thinking, and noticing that the student: (i) begins by swapping the two rightmost characters, then swapping the first two leftmost characters with the two rightmost and again swapping the two rightmost, then continuing to look for a new sequence on which to swap without repeating; (ii) shifts to the more systematic approach of fixing the first and recursively generating all swaps on the remaining three characters starting with the 11th ordering; and (iii) has missed two orderings beginning with “M” because the initial pattern did not have a systematic approach to determining the first two characters, but has been quite orderly and careful throughout.
3. Identifying that: (i) the solution is missing two orderings or that “M” is the start of only 4 orderings while the other letters have 6 and (ii) the second half of the list (for orders beginning with “L” and “W”) competently uses a powerful standard system for finding all orderings. In addition, recognizing that the student’s approach for the second half can be used to systematically list all solutions and reason that you have them all.
4. Deciding to have the student use the work in the second half of the list to reconsider or redo the work in the first half. Further, using the two missing orderings to get the student to realize the need for further work and drawing the student’s attention to the pattern in the second half of the list and the idea of using that pattern to make a complete list and reason that it is complete.

Why Defining the Construct Matters: An Examination of Teacher Knowledge Using Different Lenses on One Assessment

Chandra Hawley Orrill (corrill@umassd.edu)
University of Massachusetts Dartmouth

Allan S. Cohen (acohen@uga.edu)
University of Georgia

Abstract: What does it mean to align an assessment to the domain of interest? In this paper, we analyze teachers' performance on the Learning Mathematics for Teaching assessment of Proportional Reasoning. Using a mixture Rasch model, we analyze their performance on the entire assessment, then on two different subsets of items from the original assessment. We consider the affordances of different conceptualizations of the domain and consider the implications of the domain definition on the claims we can make about teacher performance. We use a single assessment to illustrate the differences in results that can arise based on the ways in which the domain of interest is conceptualized. Suggestions for test development are provided.

Keywords: assessment development, teacher knowledge assessment

Introduction

While test performance is generally reported as if the score assigned a participant were the goal of the assessment, the actual interest is the inferences that can be made about a learner based on that score. It is critical to ensure an assessment is measuring what it is intended to measure if such inferences are to be accurate. Thus, assessments must be written in a way that allows accurate inferences to be drawn. If we are to make claims about quantities of or changes in participants' knowledge, alignment between the content and the underlying assumptions of the domain is critical. Further, if instruction is to be impacted in positive ways by assessment data, we need to ensure that scores accurately report knowledge of the intended construct. Thus, defining the construct one is interested in measuring is vital to the assessment process.

One particularly complex domain from a measurement perspective is that of teacher knowledge. This is complex because teacher knowledge is multidimensional. The specialized knowledge teachers need for teaching (SKT) necessarily includes content knowledge, pedagogical knowledge, and understandings of how students learn (e.g., Ball, Thames, & Phelps, 2008; Baumert et al., 2010; Manizade & Mason, 2011; Shulman, 1986; Silverman & Thompson, 2008). Measuring such knowledge requires adherence to a set of beliefs about the specific construct and how it is best tested, for example, using a paper-based assessment or using feedback in a video-based open-response system (e.g.,

The Mathematics Enthusiast, ISSN 1551-3440, vol. 13, no. 1&2, pp. 93–110

2016© The Author(s) & Dept. of Mathematical Sciences-The University of Montana

Kersting, Givvin, Sotelo, & Stigler, 2010). Despite the challenges of this complex knowledge domain, if we care about whether a teacher has the knowledge necessary to support student learning, assessments need to be written to address the domain.

In the case of SKT, we are faced with not only the complexity of the domain, but also the ambiguity of what it means in practical terms for a teacher to have or exhibit particular kinds of knowledge. In our work on proportional reasoning, we have chosen to focus on how teachers understand the mathematics of proportions rather than on their pedagogical understandings related to teaching such content. However, we also acknowledge that teachers need to be able to use the content in the process of teaching, thus our interest in assessment focuses on the knowledge teachers need as it is situated in tasks that ask teachers to make sense of student thinking, analyze multiple approaches to problems, or other authentic teaching activities. Fully defining the SKT in which we are interested is outside the domain of this article. However, we rely heavily on the work of Lamon (2007) and Lobato and Ellis (2010) in our definition. For example, we know that teachers need to have the ability to conceptually connect the two values in a ratio and to understand that a third, abstractable quantity results from that connection (e.g., Lobato & Ellis, 2010). Teachers also need to understand the multiplicative relationships inherent in proportional relationships (e.g., the constant of proportionality is the multiplicative relationship of one value in the ratio to another). And, they need to understand that this corresponds to the unit rate. Further, we assert that teachers should understand how proportional reasoning connects to other areas of mathematics such as fractions and geometric similarity (Lamon, 2007; Pitta-Pantazi & Christou, 2011). We rely on research on teacher knowledge that shows that teachers struggle to use unit rate when faced with values less than one (Harel & Behr, 1995; Post, Harel, Behr, & Lesh, 1988). And, we know from a series of small studies (e.g., Riley, 2010; Son, 2010) that teachers tend to rely on cross-multiplication, which seems to obscure the breadth of knowledge they may have about proportional relationships.

Considering only the domain of proportional reasoning for teachers, one could take a number of approaches to measuring different aspects of SKT. In this study, we considered one assessment's approach to measuring the construct of proportional reasoning knowledge for teaching through the lens of our emerging definition. By undertaking this effort, we were able to further define the domain of SKT for proportional reasoning and to consider valid measurement of that domain. We relied on an approach that searched for latent classes in the data and what those latent classes might reveal about the nature of teachers' knowledge. In this setting, latent classes are statistically determined groupings of participants who shared particular aspects of patterns in their responses to the assessment items. We suspected that the data available for this study might contain latent classes given results from previous research by Izsák, Orrill, Cohen, & Brown (2010) on teachers' understanding of rational numbers. That research indicated that in a similar assessment, a two-class model fit the data better than the one-class model.

Mixture Rasch Model

For this analysis, we used the mixture Rasch model (described below) as a means of further examining the data provided through Item-Response Theory (IRT) by an assessment of teacher knowledge. The standard Rasch model is useful for tests that are designed to assess single categories of knowledge. With respect to measures of teachers' proportional reasoning, the standard model constrains the information about a teacher to a single estimate of proportional reasoning. As such, the standard model will not detect differences in the ways that examinees respond to individual items on the test. The Rasch model assumes that all examinees are drawn from a single population. When it is suspected that this may not be the case, such as when different groups of teachers have different patterns of responding to the items, then a mixture Rasch model (Rost, 1990, 1997) may be more useful.

The mixture Rasch examines patterns in responses to the assessment items that allow participants to be placed into latent classes. Latent classes are not determined a priori nor are they typically determined by more apparent commonalities such as ability, gender, or race. Members in a latent class are homogeneous on the characteristic that caused the latent class to form. Previous research with mixture IRT models in general, of which the mixture Rasch model is one example, has demonstrated that these models can address whether or not examinees exhibit the same response characteristics or whether there are groups of examinees that are latent and can only be identified by examining homogeneities in their response patterns. The groups are termed latent since they are not immediately visible simply by examining their responses. Observable characteristics like gender, height, and ethnicity are considered manifest. In contrast, characteristics such as differences in use of cognitive strategies for answering test items are considered latent until subsequent analysis can make them manifest. Previous research, for example, has found that latent classes of students that differ in their use of cognitive strategies for answering test questions can be detected (Bolt, Cohen, & Wollack, 2001; Embretson & Reise, 2000; Mislevy & Verhelst, 1990; Rost, 1990, 1997).

In a recent study of fraction knowledge that used the mixture Rasch analysis, we found that teachers in one latent class attended to the referent unit more than those in the other class (Izsák et al., 2010). These fine-grained understandings are important for mapping a domain of knowledge that teachers should have and for understanding what learning means within the domain. For example, we saw that teachers sometimes transitioned between latent classes following an intervention, albeit without exhibiting growth in their test scores. Similarly, some had significant changes in test scores without changing latent classes (e.g., Izsák, Jacobson, de Araujo, & Orrill, 2012). In our previous work, the latent classes were found to align with aspects of knowledge that are critical for meaningful understanding of fractions. We assert, therefore, that understanding latent class membership can provide important insights into the nature of teacher knowledge that ability scores alone cannot.

Methods

Sample

For the purposes of this analysis, we used the same dataset for all three analyses, but included different subsets of items for each version. While determining the latent class membership, item difficulty, and ability scores relied on the combined datasets (25 teachers from our sample plus 351 teachers from the LMT-PR sample), the analysis of the latent classes relied solely on an analysis of the 25 teachers participating in a larger study in which this work is situated. The 351 teachers in the national sample included teachers in grades 4-8 (Hill, 2008). The 25 teachers we analyzed in depth here were all teachers in grades 6-8. These 25 participants represented a convenience sample of practicing mathematics teachers drawn from across three states. They represented an array of schools including public, private, and charter schools situated in urban, suburban, and rural settings. Because the data are presented and analyzed in aggregated form throughout this study, we have not provided in-depth information about each teacher in the study.

Instrument

The instrument of interest was the Learning Mathematics for Teaching form for Proportional Reasoning (LMT-PR: Learning Mathematics for Teaching, 2007), which included 73 unique items. Consistent with all LMT assessments, the LMT-PR seeks to measure a particular construct known as mathematical knowledge for teaching (MKT; Ball, et al., 2008). MKT emphasizes the use of knowledge in and for teaching rather than focusing on the broader body of knowledge teachers may have developed (Ball et al., 2008). Items on the LMT were designed to measure common content knowledge (CCK; math knowledge used outside of teaching) and specialized content knowledge (SCK; math knowledge specifically used in teaching) (Hill, 2008). The assessment was not created using a systematic conceptualization of the domain of proportional reasoning. Rather, the development team was opportunistic, using items that captured MKT across an array of topics with some breadth and an array of mathematical tasks of teaching (Thames, personal communication). The item development team aligned items to the *Principles and Standards for School Mathematics* (NCTM, 2000). The LMT-PR form included a wide range of questions including those asking teachers to solve proportions, to select harder or easier examples for students, to explain particular proportions ideas to students, to interpret a variety of graphs, and to determine whether given situations were directly/inversely proportional. As with other LMT forms, LMT-PR included questions that asked teachers to make sense of students' work, to interpret representations, and to select items appropriate for their students. It is the need for this array of understandings that makes the construct of teacher knowledge unique and challenging to measure.

We selected the LMT-PR both because it measures the content in which we are interested and because it situates many of the tasks in the work that teachers do, as described above. Our analysis was meant to highlight the importance of aligning constructs of interest to assessments and should not be interpreted as detracting from the important role the LMT has played in the assessment of teacher knowledge. The LMT has been one of the most widely used assessments of teacher knowledge and has been used as the basis for important research showing the alignment between teacher knowledge and student performance on standardized assessments (e.g., Hill, Rowan, &

Ball, 2005). Further, the LMT paved the way for the increasingly rich discussion of the construct of teacher knowledge in mathematics.

Table 1. *Items included in each version of the LMT-PR in this analysis.*

	LMT-PR 73	LMT-PR 60	LMT-PR 54
Items Removed from full version of LMT-PR	All items included	3a-c 10 15 24a-d 32a-d	3a-c 5 8 10 12a, 12d 14b 15 18 24a-d 32a-d

As shown in Table 1, each version of the LMT used in this study was created from the full assessment. Version 1 (LMT-PR 73) for this analysis considered all 73 items in the LMT-PR¹. The second analysis (LMT-PR 60) removed 13 items that did not specifically measure proportional reasoning concepts. These items included several that asked teachers to interpret nonlinear graphs (see Figure 1 for an example). While these items featured covariational reasoning, they measured mathematics outside our construct of interest. If we wanted to carefully measure the domain to make inferences about teachers' proportional reasoning as we conceptualized it, then including items that did not specifically address proportional reasoning could create noise in the analysis.

¹ While the LMT is logically organized into sets of questions that appear to be testlets, in our past experience, those related questions have not had interdependent responses like testlets should. Participant responses on related questions appeared to be independent. Thus, it is appropriate to treat every item as a stand-alone item rather than considering them as testlets for the purposes of this analysis.

Ms. Reese and Mr. Ward celebrate student success by allowing students to eat small bags of popcorn at the end of the day. Describe what has happened over the course of the last six days in each of their classes.

Figure 1. A task stem similar to those removed to create LMT-PR 60 for this study. (Note that actual LMT items are secure.)

For the third version, which included 54 items, we worked from LMT-PR 60 and removed those items reviewers suspected would lead to errors in measurement of the proportional reasoning domain based on our interpretation of the specialized knowledge teachers need for this domain. In particular, we were concerned with items that would lead to the right answers using incorrect reasoning or incorrect answers for reasons other than limitations in understanding.

One example of such an item was the single item from the LMT-PR for which we had item response interview data from a previous study. In that study, we asked 13 middle school teachers to think aloud on an item that asked them to explain why the cross-multiplication algorithm works. In our analysis of those interviews, we found that about half of the participants who answered the item incorrectly actually understood the intended correct answer as being the most mathematically precise. However, they chose a different response driven by their interpretation of the mathematics through the eyes of a teacher, even though the question did not necessarily ask them to. That is, they chose the less accurate answer because it more precisely reflected the way they would teach students about cross multiplication. This becomes particularly relevant to the discussion of our construct because the LMT team viewed this item as working appropriately because their definition of MKT is tightly tied to the ways in which teachers use their knowledge to teach (Thames, personal communication). Thus, the item would be seen as operating appropriately from the perspective that the teachers were not applying the most appropriate understandings in their teaching situations. However, from our perspective, which is grounded in knowledge in pieces (diSessa 1988; 2006), we assert that teachers

may have a number of knowledge resources that are connected in ways that cause them to be invoked in some situations, but perhaps not in others. Working from this perspective, we interpreted items like the cross multiplication item as not working because we saw evidence that some teachers understood the mathematics of interest, but did not invoke that mathematics in expected ways in the situation presented in the item. By not using their understanding in expected ways, the participants appeared not to have the understanding, but based on their explanation, that is not an accurate assumption.

Based on our understanding and experiences working with teachers around proportional reasoning, we removed six potentially problematic items to create the 54-item version of the assessment.

Data Analysis

As mentioned above, data analysis was done using a mixture Rasch model (Rost, 1990, 1997). In the standard Rasch model (e.g., Hambleton & Swaminathan, 1985; Lord, 1980), item responses are typically scored dichotomously (e.g., 1 for a correct response and 0 for an incorrect response). The resulting data are then used to estimate parameters that describe ability for each examinee (θ_j) and difficulty for each item (b_i). Ability (θ_j) describes the amount of proportional reasoning knowledge possessed by person j . The scale for the model is centered at 0, and if $\theta_j = 0$, then person j is assumed to have the average amount of knowledge of proportional reasoning. The difficulty of the item, b_i , is expressed on the same scale as ability. The standard form of the Rasch model is given as

$$P_i(\theta_j) = \frac{\exp(\theta_j - b_i)}{1 + \exp(\theta_j - b_i)} \quad (1)$$

The item difficulty, b_i , indicates the point on the scale at which persons with ability equal to the item difficulty have a 50-50 chance of answering that item correctly. Items that are more difficult have higher difficulty parameter estimates. Similarly, items that are easier have lower difficulty parameter estimates. As an example, for item 1a (Figure 1), the item difficulty score is -2.26 for Class 1; therefore, someone whose ability is -2.26 has a 50% likelihood of answering the question correctly. In contrast, item 8 shows that a member of Class 2 to have a 50% chance of answering the item correctly, the participant would have to have an ability score of 2.736.

In our study, we used the mixture Rasch model (Rost, 1990) as estimated using a Markov chain Monte Carlo algorithm as implemented in the computer software WinBUGS (Spiegelhalter, Thomas, Best and Lunn, 2003). The mixture Rasch model can be given as

$$P_i(\theta_j) = \frac{\exp(\theta_{jg} - b_{ig})}{1 + \exp(\theta_{jg} - b_{ig})}$$

where the subscript g is used to indicate latent class. In this equation, it can be seen that ability for person j and the difficulty for item i differ depending on the latent class.

For each item, the mixture Rasch model estimated a separate item difficulty for each latent class, a separate probability of belonging in each latent class, and a separate ability estimate for each examinee. To determine the number of latent classes, we used the BIC (Bayesian Information Coefficient, Schwartz, 1973). This is a standard approach to determining fit for mixture IRT Models (Li, Cohen, Kim & Cho, 2009).

It is a well-known caution that that a statistical result does not necessarily indicate a meaningful result. In latent class analysis, for example, researchers have found that spurious latent classes may be found (Alexeev, Templin, & Cohen, 2011). In this study, the three different solutions detected in the different mixture Rasch model analyses were each composed of different test content and thus had different interpretations. Given that the mixture Rasch model was of the same family as the Rasch model used to originally calibrate the LMT items, that the latent classes detected in this study had a meaningful interpretation based on the membership of the different classes and that there were differences in performance on the content by members of each class on the different tests, it seems reasonable to infer that the latent classes were not spurious. Further, the classes we found had a clear meaning based on the membership of the classes and the differential responses of each latent class to the questions on the test. The membership of latent classes changed when the content of the test changed, which is consistent with the assumptions underlying latent class analyses. Finally, the responses offered by the members of the latent classes were helpful in determining how to characterize the different classes and the differences in performance on each of the test questions were useful in helping to characterize latent classes.

Once latent classes had been identified, we undertook analysis of the latent classes by considering those places where item difficulties varied significantly by class as well as those places where members of the lower-scoring class found items to be easier than those of the higher-scoring class. As shown in Figure 2, item difficulties were reported in standard deviation units, with easier items showing lower scores. Analyzing these items for trends in the knowledge used helps us better understand what makes the knowledge of the members of one class different from that of the other class. In our previous research (Izsák et al., 2010) we found that having item interview data substantially enhanced our ability to discriminate between classes. However, interview data were not available for the present study.

Figure 2. Plots of item difficulties for Class 1 and Class 2 in our analysis of LMT-PR. Overall, Class 2 scores are higher, but note there are items Class 1 members found easier.

In our initial analysis of the LMT-PR data from the combined dataset, we noted that the higher scoring of these classes seemed to find items focused on using cross-multiplication and other algorithms to be easier than their counterparts in the lower-scoring class. However, members of the lower-scoring class found items focused on understanding proportions and reasoning about them in a number of ways to be easier than those in the higher-scoring class. That is, the members of the higher-scoring class did better on items that allowed algorithmic thinking rather than solely focusing on reasoning about proportional relationships. These results suggested that the assessment itself may privilege teachers who are more facile with algorithms despite the field-wide emphasis on the importance of being able to reason about proportions (e.g., Lamon, 2007; Lobato & Ellis, 2010; NCTM, 2000). This finding raised interesting questions about the nature of the LMT-PR and what it might be measuring. This, in turn, led us to subsequent analyses focused primarily on 25 participants from whom we have collected completed LMT-PR forms, a prompted interview relying on a think-aloud protocol, and a face-to-face clinical interview. In the current study, we present our findings of the analysis of subsets of items on the LMT based on the performance of these 25 participants.

Results

Our aim in conducting this analysis was to consider the implications of construct definition on inferences that can be made about teachers' knowledge. Specifically, we sought to use this case as an illustration of the importance of mapping an assessment to the construct of interest. In this results section, we present the findings of our analysis of each of the three LMT-PR versions, relying on our analysis of the sample of 25 teachers. We follow with a discussion of two key issues: the impact of the definition of the domain on the overall ability scores of participants and the need to constrain assessments to

robust understandings of the mathematics of interest. In the Conclusions section, we highlight approaches that may support better assessment creation.

Ability Scores

Consistent with our position outlined at the outset, determining the construct to be measured is important for making claims. As shown in Table 2, removing the items that did not fit our definition of proportional reasoning led to a significant increase—where significance is defined as 0.3 or greater—in the scores of three participants (12%). By removing the items that did not measure proportional reasoning and those that were determined to be problematic (LMT-PR 54), seven of the 25 participants (28%) had significant increases in their ability scores. This changed the nature of our interpretations of these teachers' knowledge. This was particularly important for Bridgette, Kelly, and Kathleen who were all within one standard deviation of the mean at the outset, but ended well over one standard deviation above the mean. Interestingly, all of the significant changes in scores were increases rather than decreases.

Table 2. *Ability scores of the 25 participants on the three versions of LMT-PR. Each ability column shows the participants' abilities along a single continuum. Scores are reported as logits. All names are pseudonyms.*

	Name	Ability LMT-PR 73	Ability LMT-PR 60	Ability LMT-PR 54
1	Autumn	1.03	1.17	1.30
2	David	0.72	0.58	0.76
3	Alan	1.88	2.20	2.74
4	Ella	2.56	2.55	3.01
5	Mike	2.12	1.90	1.97
6	Bridgette	0.50	0.68	0.86
7	Allison	1.03	0.97	1.07
8	Larissa	2.12	2.20	2.51
9	Tori	0.87	0.87	0.97
10	Matt	0.87	0.77	1.07
11	Greg	2.73	2.75	2.74
12	Meagan	0.65	0.68	0.86
13	Brianna	0.15	0.22	0.36
14	Felicia	1.88	1.76	1.97
15	Eileen	1.47	1.51	1.55
16	Kelly	0.50	0.97	1.19
17	Todd	1.57	1.51	1.68
18	Kathleen	0.80	0.97	1.30
19	Patricia	1.67	1.76	1.82
20	Robyn	1.67	1.63	1.82
21	Peter	-0.84	-0.83	-0.58
22	Nancy	2.73	2.55	2.74
23	Diana	2.26	2.20	2.31
24	Christine	-0.34	-0.30	-0.48

25	Heather	2.40	2.99	3.34
----	---------	------	------	------

Table 3. *Participants' abilities for each version of the LMT-PR. The Ability scores reported here are for each latent class, but have been scaled to be comparable across latent classes.*

	Name	Mixture Ability LMT- PR 73	Class LMT- PR 73	Mixture Ability LMT- PR 60	Class LMT- PR 60	Mixture Ability LMT-PR 54	Class LMT- PR 54
1	Autumn	1.12	2.00	1.13	1.00	1.26	1.00
2	David	0.80	2.00	0.61	1.00	0.77	1.00
3	Alan	1.91	2.00	1.95	1.00	2.60	2.00
4	Ella	2.53	2.00	2.26	2.00	2.80	2.00
5	Mike	2.16	2.00	1.72	2.00	1.85	1.00
6	Bridgette	0.58	2.00	0.70	1.00	0.85	1.00
7	Allison	1.10	2.00	0.84	2.00	1.05	2.00
8	Larissa	2.15	2.00	1.96	1.00	2.40	2.00
9	Tori	0.98	2.00	0.85	1.00	0.96	1.00
10	Matt	0.96	2.00	0.77	1.00	1.05	1.00
11	Greg	2.66	2.00	2.38	1.00	2.57	2.00
12	Meagan	0.78	1.00	0.70	1.00	0.87	1.00
13	Brianna	0.34	1.00	0.30	1.00	0.43	1.00
14	Felicia	1.94	2.00	1.62	1.00	1.90	2.00
15	Eileen	1.54	2.00	1.41	1.00	1.49	1.00
16	Kelly	0.59	2.00	0.94	1.00	1.16	1.00
17	Todd	1.62	2.00	1.42	1.00	1.60	1.00
18	Kathleen	0.88	2.00	0.95	1.00	1.25	1.00
19	Patricia	1.74	2.00	1.61	1.00	1.75	2.00
20	Robyn	1.73	2.00	1.49	1.00	1.73	2.00
21	Peter	-0.56	1.00	-0.63	1.00	-0.45	1.00
22	Nancy	2.65	2.00	2.28	2.00	2.56	2.00
23	Diana	2.27	2.00	1.95	1.00	2.19	2.00
24	Christine	-0.13	1.00	-0.16	1.00	-0.34	1.00
25	Heather	2.39	1.00	2.60	2.00	3.02	2.00

Latent Class Combined with Ability Scores

When we added latent class membership to the analysis, we were able to see that these three versions yielded different results (Table 3). In essence, they measured the construct differently. First, for each of the three versions, Class 2 was the higher scoring class (Table 4). This was true despite the fact that on LMT-PR 60 80% of the participants were in Class 1 whereas in LMT-PR 73 80% of the participants were in Class 2. While we initially suspected that the computer may have switched labels in the analysis (which does happen), our analysis of the responses from the members of these classes suggested

that was not the case. In LMT-PR 73, our analysis showed that for the 25 participants, membership in Class 1 indicated more facility with ratio tables and combining ratios, which includes knowing that it is okay to add ratios to each other. Class 1 teachers also found easier those items that asked them to explain why particular relationships did or did not work as proportions (e.g., using scale factor, equivalence, division, or additive reasoning as rationales). Class 1 teachers found determining particular instances of inverse proportion to be easier than Class 2 teachers and they were better at setting up proportions for simple word problems (e.g., if two tickets cost \$5, how much do 10 tickets cost). In contrast, Class 2 was better at items that involved algorithms, including those using unit rate, scale factors, and cross multiplication. They found complex equations to be easier to set-up and verify (e.g., situations in which two people cut grass at different rates and one needs to set up an equation to determine how long it will take to mow a particular number of lawns) as well as showing more flexibility in acceptable proportion set-ups (e.g., lbs/lbs, \$/lbs, and \$/\$ are all acceptable). If the class labels had just been flipped, the groupings described for LMT-PR 73 could be reversed for LMT-PR 60, however, that was not the case. Instead, for LMT-PR 60, Class 1 seemed to be identified through items that involved explanations of why a situation is proportional, much like Class 1 for LMT-PR 73. In addition, Class 1 for LMT-PR 60 was better able to identify numbers that would make problems easier or harder for students, to identify situations in which additive reasoning was being improperly used, and to correctly interpret a variety of ways to solve proportions without a standard algorithm or equation. In contrast, Class 2 found easier those items that asked them to identify situations that were related in ways that were not proportional (e.g., linear relationships), those that asked them to model complex situations, and those that asked them to use scale factor reasoning for within measure space scale factors (e.g., given $\$3/17$ liters = $\$5/x$ liters, one can divide \$5 by \$3 to see how many times larger 5 is than 3. Then multiply that result by 17 to determine x).

Class 1 and Class 2 meant different things in these analyses despite Class 2 including 80% of the participants in LMT-PR 73 and Class 1 including 80% in LMT-PR 60. In short, while it was clear that the class separations were unique to each version, the latent classes did not clearly organize the teachers in ways that might help us make inferences beyond whether the teacher was likely to be comfortable with algorithms.

Table 4. *Mean logit scale scores for each latent class for each version of LMT-PR*

	LMT-PR 73	LMT-PR 60	LMT-PR 54
Class 1 mean	0.565 <i>n</i> =20	1.096 <i>n</i> =5	.909 <i>n</i> =14
Class 2 mean	1.595 <i>n</i> =5	1.939 <i>n</i> =20	2.233 <i>n</i> =11

More interesting was LMT-PR 54 because more teachers scored significantly higher on it and because it seemed to separate the teachers into latent classes in ways that might be more useful for measuring our construct given the more balanced distribution of

teachers between the classes. Closer examination of LMT-PR 54 class membership highlighted more mathematical sensitivity in the separation of classes. Whereas LMT-PR73 and LMT-PR 60 both had clear separations, one of the clear distinguishing features between classes was that one class found using algorithms and algebraic approaches (such as modeling equations) to be easier than the other class. In LMT-PR 54, that distinction disappeared, which suggested that the privileging of algorithmic reasoning might have been mediated in this version of the assessment. Instead, we saw finer-grained and more conceptually grounded mathematical ideas separating the classes. For example, Class 1 found items that relied on scale factors determined by the within measure space relationship to be easier, whereas Class 2 found items related to scale factors determined between measure space to be easier (e.g., given $\$3/17$ liters = $\$5/x$ liters, one can divide 17 liters by $\$3$ to determine the constant relationship, then multiply 5 by that value to answer for x). While both classes found particular items involving the combination of ratios to be easy, Class 2 seemed more able to both break down ratios, combine ratios, and to make sense of the ratio table representation. Class 2 found items that focused on why one cannot use addition to maintain equivalent ratios to be easier than Class 1. However, Class 1 was still better able to select explanations for why particular relationships did or did not work as proportions (e.g., using scale factor, equivalence, division, or additive reasoning as rationales). This was the one set of questions that Class1 consistently found easier across all versions. Class 1 in LMT-PR 54 also found the identification of easier or harder numbers (e.g., which set of numbers would make this problem harder for students to solve?) to be easier than Class 2 did. Interestingly, in LMT-PR 54, we see that Class 1 participants had an easier time with percentages as they related to proportions (for example, a sale price of 40% off is not the same as taking an additional 10% off of a 30% off price). This suggests that Class 1 might be more sensitive to questions grounded in the work of teaching (e.g., making pedagogical decisions) versus the work of solving problems.

In summary, LMT-PR 73 and LMT-PR 60 were more consistent with our initial analyses of LMT results. The class membership indicated that the higher scoring class was the one that found algorithms easier. It was not until we used the most narrowed form—the form that included only the 54 items related to proportional reasoning and removed items that seemed potentially problematic—that we were able to see differences in how participants reasoned about and with proportions and we were able to start seeing some separation tied to pedagogical concerns. This suggests that an assessment more focused on the construct of interest may yield more sensitive results.

Discussion

In this paper, we used the case of the LMT-PR to highlight the importance of aligning the construct of interest to the assessment being used. Our analysis was aided by the use of psychometric models that allow us to vary from the unidimensional analysis to which traditional IRT scores are constrained. This allowed us to look at the same participants' performance on different versions of the same assessment.

Our findings indicated that the version of the assessment that was most tailored to our definition of the construct yielded clearer information about teachers' understandings of proportional reasoning and the higher overall scores for our participants than the less focused versions. Given the emergence of high stakes testing for teacher hiring and

evaluation, being able to make clear, strong claims about teacher understanding is critical. For over one-quarter of the 25 participants in this study, scores varied significantly according to the particular items included in the assessment raising questions about what the versions of the assessment tell us about individual teacher knowledge. In contrast, at the group level (which is what the LMT is designed to measure) there was less variation in means. The group mean for the 25 participants on LMT-PR 73 was 1.39 and LMT-PR 54 was 1.49. We speculate that some of the limitation in variation is due to the overall skewing of our scores. The 25 participants we worked with were mathematically stronger than the national sample used to create the scale on which the scores were based.

In our analysis, we presented evidence that considering the assessment items' alignment to the construct of interest matters to the outcomes of that assessment and the inferences that can be made. The claims we can make about the participants changed based on the items considered, and the guidance provided by an analysis of a particular set of items could lead to potentially different implications for further learning opportunities for these participants.

Importance of Aligning Assessments to the Defined the Construct

This study demonstrated that aligning an assessment as closely as possible to the defined construct can lead to significantly different interpretations of teachers' knowledge than those assessments that are less aligned. Thus, it is important for assessments to be as well aligned to the construct as possible. The LMT-PR was created with the NCTM standards at its heart (Hill, 2008), however, those standards allowed a broad interpretation of the domain of proportional reasoning, which led to item development that spanned a variety of topics. This led to both limits in the number of items focused on key proportional ideas and the number of items focused on the mathematics to which proportions are connected. For example, there are no questions linking proportions to slope and only two questions that link proportions to similarity. There are also no questions that ask teachers about the definition of a ratio and how that definition might be the same as or different from a fraction. This is a limitation in the LMT-PR's alignment to our construct of interest.

Of course, alignment also relies on robust definitions of the construct to be measured. This is problematic in domains like teacher knowledge where the constructs are not yet well defined. It is also difficult when using existing assessments that may not fit a construct as tightly as necessary.

Our approach of constraining a pre-existing assessment to key mathematical ideas related to the construct yielded more useful information for guiding subsequent instruction and/or making claims about participants' understandings. LMT-PR 54, the version most aligned with our conception of the specialized knowledge teachers should have of mathematics, yielded information at a finer grain size than the less focused versions. In general, it seems that Class 1 found pedagogically focused questions easier (such as selecting easier or harder numbers), whereas Class 2 had more facility with manipulating equivalent ratios. This information is at a fine-enough grain size to help a professional developer plan subsequent instruction.

Additional work would need to be done to translate the findings of this study into a format that would support instructional decision-making for professional development.

But, we assert that the analysis presented here suggests there is promise in using latent class measures, combined with using focused assessments, to provide results that can provide the basis for instructional decision making.

Limitations

As with all studies, this one has a number of limitations. Here we present three major limitations. First, the national dataset included a number of 4th and 5th grade teachers for whom proportional reasoning is not in the content they teach. This may have led to very different results for the latent class analysis than if the teachers had all been from middle school. Further, our findings are limited by our lack of interview data with members of the latent classes. We could make stronger claims and, perhaps, find more similarities and differences between the latent classes were we able to hear why the teachers selected particular responses. Finally, we were limited in that the definition of the construct measured in the LMT is different from our own definition. This study sought to find the best fit between our construct definition and that LMT-PR, but that is still not as well aligned to our definition as if we had developed an assessment ourselves.

Conclusion

Clearly, the development of assessments for measuring teacher knowledge is an area in need of much more consideration (Orrill & Cohen, in press). In the domain of proportional reasoning, for example, additional work is needed to define the specific construct of interest: the knowledge teachers should reasonably be expected to know.

Test developers and users could ensure alignment between their construct of interest and the assessment using one of many available techniques that rely on mapping the assessment to the domain of interest. For example, the construct of proportional reasoning as we defined it and the assessment used may have been achieved through reliance on systematic identification of the subconstructs of interest and intentional spread of items across the subconstructs (e.g., Izsák et al., 2010).

Another approach to mapping the domain would be through the use of a Q-matrix, which provides a confirmatory approach to measuring a domain. Q-matrices, critical for cognitive diagnostic models (e.g., Izsák & Templin, in press), require the research team to identify the subconstructs to be measured and indicate which of those subconstructs each item addresses. This allows tracking of all of the subconstructs intended to be measured, thus ensuring not only that the relevant subconstructs are being measured but also that they are being paired in multiple ways so that one idea does not obstruct the other. For example, if we are interested in the use of representations and the teachers' understandings of combining ratios, we would not want all of the combining ratios tasks to include ratio tables because that representation may be unfamiliar to teachers, thus masking the participants' actual knowledge of combining ratios.

In the end, careful consideration of the alignment of the assessment to the construct is important for making claims of validity. Given the widespread move toward using assessments of teacher knowledge for high stakes decision-making, this becomes even more important. The study presented here shows that the same teachers have the perception of more knowledge or less simply based on the items included in the analysis.

Acknowledgements

Work on this paper was supported by the National Science Foundation through grant number DRL 1054170. The opinions expressed here are those of the authors and do not necessarily reflect the views of the NSF. The authors wish to thank Heather Hill for sharing the LMT-PR data with us as well as Mark Thames for his conversations about MKT. We also wish to thank Eric Gold, Dave Kamin, Tim Marum, Rob Nanna, Dennis Robinson, Ryan Robidoux, and Kaitlyn Walsh Rodrigues for their assistance in the initial analysis of the LMT-PR assessment.

Works Cited

- Alexeev, N., Templin, J., & Cohen, A. (2011). Spurious latent classes in the mixture Rasch model. *Journal of Educational Measurement*, 48, 313–332.
- Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, 59, 389–407.
- Baumert, J., Kunter, M., Bum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., ..., & Tsai, Y. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, 47(1), 133–180.
- Bolt, D. M., Cohen, A. S., & Wollack, J. A. (2001). A mixture item response for multiple-choice data. *Journal of Educational and Behavioral Statistics*, 26, 381–409.
- diSessa, A. A. (1988). Knowledge in pieces. In G. Forman & P. Pufall (Eds.), *Constructivism in the computer age* (pp. 49–70). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- diSessa, A. A. (2006). A history of conceptual change research: Threads and fault lines. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 265–282). New York: Cambridge University Press.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum.
- Hambleton, & Swaminathan, H. (1985). *Item response theory: Principles and application*. Boston, MA: Kluwer-Nijhoff Publishing.
- Harel, G., & Behr, M. (1995) Teachers' solutions for multiplicative problems. *Hiroshima Journal of Mathematics Education*, 3, 31–51.
- Hill, H. C. (2008). *Technical report on 2007 proportional reasoning pilot Mathematical Knowledge for Teaching (MKT) measures learning mathematics for teaching*. Ann Arbor, MI: University of Michigan.
- Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, 42(2), 371–406.

- Izsák, A., Orrill, C. H., Cohen, A., & Brown, R. E. (2010). Measuring middle grades teachers' understanding of rational numbers with the mixture Rasch model. *Elementary School Journal*, 110(3), 279–300.
- Izsák, A., Jacobson, E., de Araujo, Z., & Orrill, C. H. (2012). Measuring growth in mathematical knowledge for teaching fractions with drawn quantities. *Journal for Research in Mathematics Education*, 43(4), 391–427.
- Izsák, A., & Templin, J. (in press). Coordinating descriptions of mathematical knowledge with psychometric models: Opportunities and challenges. In A. Izsák, J. T. Remillard, & J. Templin (Eds.), *Psychometric methods in mathematics education: Opportunities, challenges, and interdisciplinary collaborations* (pp. xxx–xxx). Journal for Research in Mathematics Education monograph series. Reston, VA: National Council of Teachers of Mathematics.
- Kersting, N. B., Givvin, K. B., Sotelo, F. L., & Stigler, J. W. (2010). Using video to predict student learning of mathematics: Further explorations of a novel measure of teacher knowledge. *Journal of Teacher Education*, 61(1–2), 172–181.
- Lamon, S. J. (2007). Rational numbers and proportional reasoning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 629–667). Charlotte, NC: Information Age Press.
- Learning Mathematics for Teaching (2007). *Survey of teachers of mathematics: Form LMT PR-2007*. Ann Arbor, MI: University of Michigan.
- Li, F., Cohen, A.S., Kim, S.-H., & Cho, S.-J. (2009). Model selection methods for dichotomous mixture IRT models. *Applied Psychological Measurement*, 33(5), 353–373.
- Lobato, J., & Ellis, A. B. (2010). *Essential understandings: Ratios, proportions, and proportional reasoning*. In R. M. Zbieck (Series Ed.), *Essential understandings*. Reston, VA: National Council of Teachers of Mathematics.
- Lord, F. N. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Erlbaum.
- Manizade, A. G., & Mason, M. M. (2011). Using Delphi methodology to design assessments of teachers' pedagogical content knowledge. *Educational Studies in Mathematics*, 76, 1883–207. DOI: 10.1007/s10649-010-9276-z
- Mislevy, R. J., & Verhelst, N. (1990). Modeling item responses when different subjects employ different solution strategies. *Psychometrika*, 55, 195–215.
- National Council of Teachers of Mathematics (NCTM) (2000). *Principles and standards for school mathematics*. Reston, VA: Author.
- National Governors Association & Council of Chief State School Officers (NGA & CCSSO) (2010). *Common core state standards mathematics*. Washington, DC: Author.
- Orrill, C. H., & Cohen, A. (in press). Purpose and conceptualization: Examining assessment development questions through analysis of measures of teacher

- knowledge. To appear in *Journal for Research in Mathematics Education Monograph*.
- Pitta-Pantazi, D., & Chritou, C. (2011). The structure of prospective kindergarten teachers' proportional reasoning. *Journal of Mathematics Teacher Education*, 14(2), 149–169.
- Post, T., Harel, G., Behr, M., & Lesh, R. (1988). Intermediate teachers' knowledge of rational number concepts. In Fennema, et al. (Eds.), *Papers from First Wisconsin Symposium for Research on Teaching and Learning Mathematics* (pp. 194–219). Madison, WI: Wisconsin Center for Education Research.
- Riley, K. R. (2010). Teachers' understanding of proportional reasoning. In P. Brosnan, D. B. Erchick, & L. Flevares (Eds.), *Proceedings of the 32nd annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 1055–1061). Columbus, OH: The Ohio State University.
- Rost, J. (1990). Rasch models in latent classes: An integration of two approaches to item analysis. *Applied Psychological Measurement*, 14, 271–282.
- Rost, J. (1997). Logistic mixture models. In W.J. van der Linden & R.K. Hambleton (Eds.), *Handbook of modern item response theory* (pp. 449–463). New York: Springer.
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6, 461–464.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14.
- Silverman, J., & Thompson, P. W. (2008). Toward a framework for the development of mathematical knowledge for teaching. *Journal of Mathematics Teacher Education*, 11, 499–511. DOI: 0857-008-9089-5
- Son, J. (2010). Ratio and proportion: How prospective teachers respond to student errors in similar rectangles. In P. Brosnan, D. B. Erchick, & L. Flevares (Eds.), *Proceedings of the 32nd annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 243–251). Columbus, OH: The Ohio State University.
- Spielgelhalter, D., Thomas, A., Best, N., & Lunn, D. (2007). *WinBUGS with DoodleBUGS*. Medical Research Council, Imperial College and MRC, UK.

Knowledge for Equitable Mathematics Teaching: The Case of Latino ELLs in U.S. Schools

Aaron T. Wilson (aaron.wilson@utrgv.edu)¹
University of Texas Rio Grande Valley, USA

Abstract: This paper reports the exploration of an aspect of knowledge needed for equitable mathematics teaching. Pedagogical Content Knowledge for Teaching Mathematics to English Language Learners (PCK-MELL) was proposed as a theoretical knowledge construct, a subdomain of MKT, and the construct was investigated through a process of survey instrument development and administration. The survey contained items intended to measure teachers' knowledge of the obstacles encountered by ELLs in math classes, of the resources that ELLs draw upon, and of instructional strategies for teaching ELLs. Analysis of middle school mathematics teachers' responses ($N = 42$) offered insights into how to improve the reliability and measurement validity of this sort of instrument, as well as directions for further theory development.

Key words: English Language Learners; mathematical knowledge for teaching; deficits; affordances

Introduction

Many mathematics teachers worldwide are finding new languages and new cultures in their classrooms. Since the 1970s the number of students who are English Language Learners (ELLs) in United States schools has grown and continues to grow dramatically (U.S. Department of Education, 2012; Payán & Nettles, 2008; Francis et al., 2006; Capps et al., 2005). Mathematics education researchers and others have observed, based upon the lower relative performance of these students on standardized exams of mathematics achievement in comparison with their mainstream (non-ELL) counterparts, that many of these ELLs have infrequently had equitable opportunity to learn mathematics in U.S. schools and have often been underserved by their teachers and schools (Center on Education Policy, 2010; Abedi & Herman, 2010). Furthermore, researchers have noted that using measurements of mathematical achievement to draw attention to disparities between student groups in fact offers little, if any, contribution toward promoting equity in education (Gutiérrez, 2008), and that focusing on achievement gaps may only perpetuate negative stereotypes. Yet, measurements of mathematical achievement have also shown that some teachers and school districts appear to serve their ELLs better than do others.

¹ A portion of this research was conducted in pursuit of a doctoral dissertation under the direction of M. Alejandra Sorto, Texas State University-San Marcos. The research was funded in part by the National Science Foundation Grant 1055067, M. Alejandra Sorto, principal investigator. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation. The author would also like to express appreciation to Drs. Mark Hoover and Reidar Mosvold as well as to the other anonymous reviewers of *The Mathematics Enthusiast* for the thoughtful feedback given.

For example, consider the following phenomenon that occurred in a large southwestern state in which approximately 16% of all public school students are ELLs². The Samsonville School District and the Wilkins School District (pseudonyms) both had significant percentages of ELL students, 25% and 33% of their approximately 80,000 and 50,000 respective student populations. Samsonville ELLs were approximately 85% Latino students, i.e., students from Spanish-speaking backgrounds, while Wilkins' ELLs were 99% Latinos. Sixty percent of Samsonville students were classified as economically disadvantaged compared to 95% of Wilkins students. Furthermore, less than 40% of Samsonville teachers were Latinos, while in the Wilkins district greater than 70% of teachers were Latinos. Although the characteristics of the students who were ELLs in these two school districts were largely equivalent in terms of language background and ethnicity, over several years Wilkins ELLs performed significantly higher than Samsonville ELLs on the mathematics portion of the state standardized test. To what could this differential performance be attributed?

In trying to explain the above phenomenon, one may ask the obvious question of whether the teachers in the Wilkins district do something differently than do the Samsonville teachers which results in the differential mathematics achievement among ELLs. One may also ask the question of whether there is something that the Wilkins teachers know about teaching mathematics to ELLs that capacitates them to more effectively instruct ELLs and that is less well known among Samsonville teachers. That is, one may seek to explain this phenomenon from the perspective of instructional practice or from the perspective of the teachers' knowledge that informs their practices.

There are also several reasons why it may be valuable to investigate a phenomenon such as this from the later perspective, that of teachers' knowledge. Researchers concerned with the assessment of ELLs have asserted that "with the rapid growth of ELL populations states should place a substantial focus on increasing teacher knowledge of current ELL issues...including pre-service teacher education and continuing teacher education" (Wolf, Herman, & Dietel, 2010, pp. 8–9). Hence, successfully characterizing knowledge that promotes achievement among ELLs would add to theories of mathematics teachers' knowledge. It would also fill a void of content for educator textbooks and professional development materials useful for equipping teachers for the work of teaching mathematics to ELLs (Watson et al., 2005). Ultimately and importantly, it would more fully complete the picture of essential elements that inform equitable mathematics teaching.

The research reported in this paper was done as part of a larger study that attempted to explain the difference in ELL achievement seen in the Samsonville and Wilkins school districts from the viewpoint of teacher qualities. That larger study looked at a number of important teacher and student variables in the hope of identifying the teacher qualities and instructional moves that resulted in higher mathematics achievement for ELLs. One among the many teacher variables to be measured in as a possible predictor of ELLs' achievement was teachers' mathematical knowledge for teaching, MKT (Ball, Thames, & Phelps, 2008). Yet, because of the

² This account regarding the Samsonville and Wilkins school districts arose through conversations between school district administrators and university researchers. Reporting of state standardized test results on the website of the state department of education website between 2002 and 2009 had revealed the higher performance of Wilkins ELLs on the math portion of the state standardized test. Administrators in some districts around the state and nation expressed interest in knowing what Wilkins teachers were "doing" with their ELLs. The study that resulted in part from those conversations was funded by the National Science Foundation (DRL-1055067).

special cultural and linguistic context of that study, the researchers sought an instrument that could capture teachers' MKT that informed their work with students who came to school with languages and prior mathematics learning that might vary from English and from traditional mathematical algorithms and notations used in U.S. schools. However, at the time of the study there were—and still are today—several limitations to doing an investigation of this sort, i.e., a study that can describe and measure the kind of mathematical knowledge needed by teachers of students who are still learning the language of instruction, in this case ELLs. The foremost of these is that, while there exist abundant strategies for teaching mathematics to ELLs, there is a shortage in theory about what effective math teachers of ELLs really need to know. Because of this lack of theory there is also a lack of research tools for investigating this knowledge. For example, most current instruments that measure mathematics teachers' knowledge – and the Learning Mathematics for Teaching (LMT) measures in particular (Hill & Ball, 2004) – do not address ELLs in particular, or aspects related to equity in general, and as a result, may fail to capture many aspects of knowledge for equitable teaching.

This paper narrates an initial attempt to create a research instrument capable of measuring knowledge for teaching mathematics as it is used by teachers in linguistically and culturally diverse classrooms containing large number of ELLs. The work reported here is fundamental in the sense that it was done in hopes of “laying the bearings” for observing MKT in linguistically and culturally diverse contexts and of building the capacity to study this aspect of knowledge for equitable mathematics teaching. Specifically, the goal of this study was to create a survey instrument that could be used to capture the particular kind of knowledge that mathematics teachers like those in the Wilkins district called upon when teaching ELLs. But how could that kind of mathematical knowledge be defined? To what extent could a survey be designed that measures that knowledge reliably? And further, to what extent does it even seem valid to define knowledge for teaching mathematics to ELLs as a special kind of knowledge? Is this a valid construct? Answers to these important questions would determine the usefulness of the survey to be created. Responses to questions like these are given here in the hope that the interested reader may appreciate the nuanced nature of the knowledge being addressed and in hopes that the foundational attempt described in this paper to measure that knowledge may only open the way for even more illuminating work along these lines.

To summarize, if we are to understand the differential performance that some mathematics teachers have with ELLs, like that described above, and to use this understanding to promote equitable access to mathematics instruction for all students, then there is the need to better understand the particular role that teachers' knowledge plays in the context of teaching math to ELLs. Thus, there is an initial need to first develop and test theory related to mathematics teachers' knowledge for teaching ELLs and to then develop research tools capable of observing and describing such knowledge. The ultimate goal of the study reported here was to produce a viable research instrument for measuring teachers' knowledge for teaching mathematics to ELLs and this process was guided by theory in survey and scale development (DeVellis, 2003; Dillman, Smyth, & Christian, 2009; Schuman & Presser, 1996). As a result, this work followed a somewhat linear fashion that commenced with a review of literature with the purpose of defining a theoretical construct, that is, defining what knowledge is needed by mathematics teachers of ELLs. Following this a test blueprint was outlined and original survey items were developed according to the knowledge domains defined in the blueprint. Finally responses were obtained from math teachers for the main purpose of evaluating the reliability and validity of the measure. The sections of this paper that follow elucidate the path that was

taken in trying to capture this knowledge, showing in detail how knowledge for teaching ELLs mathematics was hypothesized based upon the research literature and classroom observations conducted, how the instrument was developed and administered, and finally what was learned about this knowledge as a result of administering the survey.

Conceptualizing Mathematical Knowledge for Teaching ELLs and Developing a Measure

Consulting the Literature: Connecting MKT to ELLs

As an entry into investigating the kind of knowledge needed for equitable mathematics teaching generally, by looking at mathematics teachers of ELLs specifically, this study began by examining the intersection between two existing strands of inquiry: research concerning mathematics teachers' knowledge and research concerning ELLs in the mathematics classroom. Highlights from some of the essential theory and findings that guided this study follow.

An important component of math teachers' knowledge is profound mathematical content knowledge (Ma, 1999, for example) and this kind of deep mathematical knowledge was assumed here to be important for math teachers of ELLs as well. Furthermore, it was assumed that, like all teachers, teachers of ELLs have *knowledge-in-practice* (Schon, 1983), i.e., knowledge that is gained by their practice of teaching ELLs and from their instructional experiences related to the particular linguistic and cultural background that their ELLs bring to the classroom. Shulman (1986) connected content knowledge to this kind of practice-based knowledge by explaining how teachers' knowledge of the content that they teach is shaped by the pedagogy that they practice. To Shulman (1986), *pedagogical content knowledge* (PCK) "includes an understanding of what makes the learning of specific topics easy or difficult" as well as "the most useful forms of representation of those ideas...—in a word, the ways of representing and formulating the subject that make it comprehensible to others" (p. 9). The above utterances concerning teachers' PCK, which have admittedly been reprinted in many a literature review because of their lucidity, are repeated here because of the role that the terms *easy*, *difficult*, and *representations* played in determining a connection between the two strands of research central to the present study and in informing survey development, as is explained in a later paragraph.

Recent research concerning mathematical knowledge for teaching, MKT, has given a robust explication of Shulman's (1986) PCK in the context of mathematics teaching (Ball, Thames, & Phelps, 2008; Hill, Ball, & Schilling, 2008; Hill, Blunk, Charalambous, Lewis, Phelps, Sleep, & Ball, 2008; Hill, Rowan, & Ball, 2005). These researchers proposed the existence of a system of related knowledge constructs, some more mathematical and some more pedagogical, that together constitute MKT. Furthermore, in their Learning Mathematics for Teaching (LMT) project, they have investigated this knowledge through the development of psychometric instruments, pen-and-paper (survey) tools, for measuring the knowledge in these domains (Hill, Schilling, & Ball, 2004).

Two subdomains of MKT, as theorized by the above researchers, were identified as particularly relevant to this study: knowledge of content and students (KCS) and knowledge of content and teaching (KCT). These researchers define KCS as "knowledge of how students think about, know, or learn this particular content" (Hill, Ball, & Schilling, 2008, p. 375). Furthermore, they define KCT as "mathematical knowledge of the design of instruction, [including] how to choose examples and representations, and how to guide student discussions toward accurate mathematical ideas" (Hill, Ball, Sleep, & Lewis, 2007). Within this model of PCK, the two

domains of KCS and KCT seemed to be of special interest to the present study because they include awareness of students' background and awareness of instructional decisions appropriate to the teaching of specific mathematics topics. The particular linguistic and cultural qualities that ELLs bring to the learning of mathematics, which qualities often vary from those found in non-ELLs, seemed to have potentially the most practical impact on these two domains of teachers' PCK. Knowledge of mathematics and of ELLs and of appropriate instructional decisions that can be made on their behalf seemed central to the kind of knowledge under investigation by this study.

Moschkovich (2002) has investigated ELLs in the mathematics classroom extensively and was among the first mathematics education researchers to begin to make a shift in the perspective taken when studying ELLs. She took a sociocultural view of ELLs as mathematics learners and observed that, historically, much research concerning ELLs had taken a deficit perspective which focuses the research lens on the obstacles, sources of *difficulty*, that ELLs encounter in mathematics classrooms. Because of this lens, such studies, she noted, had failed to observe the resources that ELLs often draw upon to make *easier* the learning and expressing of mathematics. Factors that have been observed to cause difficulty for ELLs in mathematics classes are many, including, for example: language of instruction and limited English proficiency (Cuevas, 1984), word problems and linguistic complexity (Llabre & Cuevas, 1983; Martiniello, 2009), polysemy (Lager, 2006), and whole-class, teacher-centered instruction format (Chang, 2008), to name a few. Conversely, much research since the 1990s has taken note of factors that serve ELLs as resources upon which they can draw to do and express valid mathematics—even if at times using grammatically invalid English, or even no English at all. Such factors include: gesturing (Shein, 2012), first language and bilingualism (Gutiérrez, 2002; Khisty & Morales, 2004; Moschkovich, 2002; Sorto, Mejía Colindres, & Wilson, 2014), non-linguistic mathematical representations (Martiniello, 2009), and prior and cultural knowledge (Gutiérrez, 2002; Gutstein, Lipman, Hernandez, & Reyes, 1997; Henderson & Landesman, 1995). Studies mentioned in this paragraph represent ways in which researchers have investigated the obstacles that ELLs face in mathematics classrooms and have also recently begun to perceive the strengths that many ELLs possess for learning mathematics.

Here a connection can be made between mathematics teachers' knowledge and ELLs. Moschkovich (2002) saw a dichotomy in perspectives taken by mathematics education researchers concerned with ELL issues; they took either a deficit or an affordance perspective. And as has been mentioned above, Shulman (1986) also saw a dichotomy; he saw that teachers' knowledge includes an understanding of the things that make learning *easy* or *difficult*, as mentioned earlier. These two sets of extremes—deficits versus affordances and difficult versus easy—constitute the link found in the present study between the two bodies of research concerning mathematics teachers' knowledge and ELLs. More precisely, the two divergent perspectives concerning ELLs of deficits and affordances, with their different research perspectives, were taken as potentially illuminating the very domains of PCK—KCS and KCT—central to teaching mathematics to ELLs. It seemed that experienced mathematics teachers of ELLs should potentially have pedagogical content knowledge by which they perceive both the *obstacles* (deficits) that their ELLs face and also the *resources* (affordances) that their ELLs draw upon in mathematics classes. These two divergent research perspectives (and their respective results) informed the method used in this study significantly, as will now be seen.

Developing the PCK-MELL Survey

In developing the survey in this study it was hoped that the understanding of knowledge needed by mathematics teachers of ELLs as well as the survey itself could be closely tied to current conceptions of mathematics teachers' knowledge, with a view to possibly investigating the relationship between this construct and related constructs in future studies. As a result, the theoretical basis of the survey was conceived of in connection with the theory behind MKT and the items themselves were written in a similar fashion as the multiple-choice LMT measures (Hill & Ball, 2004), although the measurement purpose and the mathematical topics and contexts were all novel. The process of survey development became a means of exploring this aspect of mathematics teachers' knowledge.

Pedagogical Content Knowledge for Teaching Mathematics to ELLs

Drawing from a review of the literature, a sample of which is given above, a hypothetical framework of pedagogical content knowledge for teaching mathematics to ELLs (PCK-MELL) was first developed (Figure 1).

Figure 1. Hypothetical framework of PCK-MELL, proposed here as a subset of MKT (Hill, Ball, & Schilling, 2008).

The framework in Figure 1 makes several structural hypotheses regarding PCK-MELL. Firstly, it identifies the construct as a subset of pedagogical content knowledge and places it squarely within the larger framework of MKT (Hill, Ball, & Schilling, 2008). Secondly, it further embeds PCK-MELL within this framework by identifying it with two specific domains of MKT: knowledge of content and students (KCS) and knowledge of content and teaching (KCT). Thirdly, it posits three subdomains of PCK-MELL: knowledge of obstacles often encountered by

ELLs in math classes (OBST), knowledge of resources that ELLs draw upon when learning math (RESRC), and knowledge of instructional strategies for usage with ELLs in math classes (STRAT). Although not referenced in the brief summary of literature above, knowledge of instructional *strategies* was included because it seemed central to both Shulman's (1986) understanding of PCK and to Ball, Thames, & Phelps' (2008) theorization of MKT (specifically the KCT domain), and because a considerable number of strategies have been posited for teaching ELLs (see for example, Coggins, Kravin, Coates, & Carroll, 2007; Echevarría, Short, & Vogt, 2007). Finally, the arrows between the OBST and RESRC domains and between these and the STRAT domain above suggests a theoretical relationship between these three subdomains: knowledge of obstacles may be related to knowledge of resources, and knowledge in both of these subdomains may inform teachers' knowledge of, and especially selection of, strategies. (It may not be at once clear to the reader why the OBST and RESRC domains were suspected to be related; indeed, the affordance and deficit perspectives of researchers looking at ELL students seem to contradict each other significantly. This suspicion arose through reflection on the conceptual process of writing items in both of these domains and the grounds for this suspicion will be clearly explained in the presentation of the OBST survey item below.)

Furthermore, the following testing framework (Table 1) was crafted to serve as a guide to survey item development. The testing framework again delineates the three proposed knowledge domains of OBST, RESRC, and STRAT and then associates with these domains a number of specific aspects of knowledge within each domain. All of the aspects provided in this framework were derived specifically from research results. That is, in so far as it seemed possible to classify select findings from the literature as either identifying and explaining particular obstacles encountered by ELLs (such as linguistic complexity, Martiniello, 2009), or as describing resources that ELLs draw upon (such as their bilingualism, Moschkovich, 2002), or as enumerating instructional strategies for usage with ELLs (see for example, Coggins, Kravin, Coates, & Carroll, 2007), then those concepts from the literature concerning the mathematics education of ELLs were placed in the framework and hypothesized to be elements of PCK-MELL about which experienced teachers of ELLs should be familiar. At least three assumptions are being made here: first, that the research findings concerning ELLs accurately depict elements of the 'real' situation concerning many ELLs in mathematics classes, second, that the findings are generalizable in contexts beyond those in which the research occurred, and third, that teachers who have PCK-MELL will have gained, possibly through formal training but much more likely through actual teaching experience, an understanding of at least some of those very same elements described in the research literature and identified in the testing framework. In essence the reasoning behind the usage of the testing framework was as follows: if research such as Martiniello's (2009) finds that the linguistic complexity of word problems is an explanative factor in the differential (lower) performance of ELLs in mathematics, then not only is the difficulty caused by linguistic complexity something about which teachers of ELLs *should be aware*, but it even seems probable that experienced teachers of ELLs *will actually be aware* of ways in which linguistically complex word problems can cause difficulty for their students. The survey items were then written for the purpose of capturing this awareness (or lack of awareness). Assumptions like these may be dangerous. Yet, some assumptions had to be made since this effort was an initial attempt at hypothesizing about the very nature and contents of the knowledge domains in question. It was hoped that results from the survey administration would either validate, or indeed invalidate and lead to improvements of, the survey.

Two final notes regarding the testing framework are in order. First, the aspects given in the STRAT domain are merely an adaption of Chval & Chávez’s (2011) synthesis of research-based strategies for instructing ELLs in mathematics. Many specific strategies for teaching ELLs have been posited and their list seemed to be general enough that specific tools such as “word walls”, posters and manipulatives, for example, could be taken as mere instances of the broader aspects of strategic knowledge in the list. Finally and most importantly, it must be made clear that this testing framework is not now and was not believed at the time of the study to be exhaustive. There are, no doubt, many more aspects of the knowledge domains that are not represented here. The items in the testing framework are a best attempt given limited time and resources.

Table 1. *PCK-MELL Testing Framework*

DOMAIN	ASPECTS
OBST Knowledge of Obstacles encountered by ELLs in mathematics classes	1. Limited English proficiency in speaking, reading and writing
	2. Word problems <ul style="list-style-type: none"> a. Specific words (vocabulary), multiplicity of words (linguistic complexity), shifts of application, polysemy
	3. Classroom format <ul style="list-style-type: none"> a. High speech formats (direct teaching versus indirect or peer-based methods)
	4. Assessments <ul style="list-style-type: none"> a. Low performance because of: <ul style="list-style-type: none"> i. Word problems, time limitation, high stakes, cultural-irrelevance
RESRC Knowledge of Resources that ELLs draw upon in learning mathematics	1. Linguistic creativity, mathematics discursive and communicative ability <ul style="list-style-type: none"> a. Fluency in L₁ (i.e., first language) b. Bilingualism c. Usage of gestures, objects, and verbal <i>inventions</i> to convey meaning
	2. Linguistic and cultural identity <ul style="list-style-type: none"> a. Personal association with cultural icons, people, values, traditions, etc. b. Appreciation of first language
	3. Prior mathematical knowledge including knowledge of and fluency with alternative or “foreign” mathematical notations and algorithms
STRAT Knowledge of instructional strategies that help ELLs in mathematics	1. Usage of students’ background knowledge—academic, linguistic and cultural—to promote understanding
	2. Maintenance of classroom environment rich in linguistic and mathematics content
	3. Emphasis on meanings of words and/or provisions for students’ usage of multiple modes of communication to express mathematics
	4. Usage of visual supports to—gestures, objects, illustrations—to convey the meanings of classroom conversations
	5. Connection of mathematical language with multiple forms of mathematical representation
	6. Available visual display of classroom mathematics concepts, representations and words during instruction
	7. Rich usage of students’ own mathematical writings and speech with opportunity for them to make revisions

Survey Development

To inform and augment the development of the PCK-MELL theoretical and testing frameworks, more than thirty hours of original classroom observations were also conducted in middle school (and a few high school) mathematics classrooms composed of large numbers of ELLs with a view to 1) situating the development of survey items within authentic and mathematically specific teaching instances while 2) populating the testing framework with both research-based and classroom-observed aspects of the three knowledge domains. Then, based upon the testing framework and observations, a large number of survey items were developed, from which an approximately equivalent number of items across the three domains of OBST, RESRC, and STRAT were retained for the final 32-item instrument after a pilot version had been sent for review and comment to a number of mathematics educators and experts in ELL and teacher knowledge issues at different institutions in the United States. To give a sense of the final instrument which was the basis of the statistical analysis that follows, three exemplary survey items are given here.

The PCK-MELL survey item presented below was derived from a classroom observation in which a math teacher taught a geometry lesson to 9th and 10th grade students who were recent immigrants to the United States, more than 95% of which were Latinos. Both the mathematical situation and the answer options are authentic in the sense that the teaching situation and words used actually happened in a mathematics classroom. Furthermore, the mathematical topics and teaching tasks represented in this item are typical of the work of mathematics teaching. The “correct” answer to the item, as with all items in the survey, was based upon theory found in the research literature concerning effective (or ineffective) practice for teaching math to ELLs. The item below was designed to measure teachers’ knowledge of the linguistic obstacles (OBST) that ELLs may encounter in mathematics classes.

Mrs. Tash's students, many of whom are Latino English Language Learners, have been working in groups to discover how surface area changes when the dimensions of an object are changed by a scale factor. On the previous class day, students used rulers and a rectangular prism to complete the table below.

Picture of Solid	Name of Solid	Dimensions	Scale Factor	New Dimensions	New Surface Area, SA	Conclusions
			k = 2			

Which word in the table may cause the most trouble for the English Language Learners in the classroom?

- A. Solid
- B. Dimensions
- C. Conclusion
- D. Surface

Figure 2. Sample PCK-MELL survey item, OBST domain.

The item in Figure 2 presents several English mathematical words that could be unknown to English Language Learners. Among these four words, three of them (solid, dimensions, and

conclusion) have direct Spanish-English cognates (*sólido*, *dimensiones*, and *conclusión*) and may be more readily known to Latino ELLs as a result. Furthermore, in classrooms of Latino ELLs teachers' knowledge of the Spanish language may be helpful in answering the particular item above. To control for this effect knowledge of the Spanish language was assessed as a covariate on the survey as well. Yet it is also not difficult to imagine that even the observant, but non-Spanish-speaking teacher of ELLs may perceive *through teaching experience alone* both the English words that cause difficulty for ELLs and the words that ELLs can use to leverage their understanding. That is, non-Spanish-speaking teachers may also gain through experience an awareness of the English words that their Spanish-speaking ELLs stumble on or readily comprehend.

Furthermore, this item and the theory behind its hypothetically "correct" answer can be used to show why knowledge of *obstacles* encountered by ELLs in math was hypothesized to be related to knowledge of their *resources*: while language of instruction and linguistic complexity can cause difficulty for ELLs on one hand (Cuevas, 1984; Llabre & Cuevas, 1983; Martiniello, 2009), the ELLs' first language and their bilingualism can also give them access to the same mathematical ideas on the other (Gutiérrez, 2002; Khisty & Morales, 2004; Moschkovich, 2002; Sorto, Mejía Colindres, & Wilson, 2014). Teachers' knowledge of both of these functions of language in mathematics classrooms may be related. As a result, teachers' knowledge of the obstacles to and resources for ELLs' learning of mathematics may also be related.

The following figure represents a survey item in the RESRC domain.

To solve this problem, students had to divide 1836 by 36. Sara and Josué used these two different methods.

Sara's Method

Josué's Method

Who's division method is correct?

- A. Both methods are correct
- B. Josué's
- C. Sara's

Figure 3. Sample PCK-MELL survey item, RESRC domain.

The survey item above presents to two slightly different notations of the long division algorithm and then requires the respondent to select the correct method for performing the long division. Sara's method is equivalent to the standard algorithm traditionally used and taught in schools in the United States while Josue's method is an equivalent algorithm, but that uses notation commonly found in Central America. This item was inspired by conversations with

Latino immigrants to the United States for whom this notation had been the standard for long division. The item was included in the PCK-MELL survey as a potential indicator of teachers' practical knowledge of resources that some ELLs (especially Latino immigrant students) may bring to the math classroom, namely, their prior mathematical knowledge, including valid, yet perhaps alternative, mathematical algorithms and notation.

The final item below was intended to serve as a STRAT item, measuring teachers' knowledge of effective instructional strategies for teaching ELLs.

Mr. Garza asks his class a question about the contents of the grape-limeade drink below: "How do you find out what proportion the grape juice is of the drink?"

Ciarra, an English Language Learner, responds: "Cuenta los grises y lo divides entre cien."
Which is the BEST way for Mr. Garza to respond?

- A. He should ask Ciarra to rephrase her answer using English
- B. He should acknowledge that Ciarra tried to answer his question, but then ask another student to respond in English
- C. He should acknowledge that Ciarra tried to answer his question, but then explain the correct answer in English so that all of the students understand
- D. He should ask Ciarra to give her response again, but to use the pictured image and gestures to show what she means

Figure 4. Sample PCK-MELL survey item, STRAT domain.

The survey item above, taken from a situation that occurred during the observation of a 7th grade mathematics classroom having many ELLs, involves a strategic decision on the part of the teacher: how should the teacher respond to the student who has answered the English mathematics question using the Spanish language? (And does the fact that the student actually gave a concise and mathematically valid answer affect teachers' responses to the item? Again, responses to this item may also be informed by the teachers' knowledge or ignorance of the Spanish language and this effect was controlled for in asking respondents concerning their level of knowledge of that language.) Originally, this item was intended to serve as an indicator of teachers' knowledge of mathematics instructional strategies for teaching ELLs, the STRAT domain. Yet, as the theoretically "correct" answer to this question was actually option "D", then selecting the correct also implied knowledge that ELLs' first language was a resource in mathematics classes rather than a deficit. Not surprisingly then the item was later found to correlate more strongly with other items in the RESCRC domain. Furthermore, the intersection (overlapping) of knowledge domains required by items like these hints at a particular weakness of this initial exploratory effort at defining the construct: in hindsight, the domains were not

altogether well-defined and items such as this, which may serve as indicators of knowledge in more than one domain, reflect the need for even more careful theory development and rigorous item-writing.

Evaluating the Instrument

Following the theoretical development and writing of survey items, the intention in this study was to obtain enough responses in order to be able to evaluate the extent to which the instrument developed above measured the knowledge that it indeed was designed to measure. The survey instrument described above first underwent pilot study using pre-service mathematics teachers ($N = 142$) at a major university in a southwestern U.S. state having large numbers of ELLs. Following this, responses to the final instrument were obtained via internet survey from forty-two ($N = 42$) in-service middle school mathematics teachers around the state. Thirty-one of the teachers were female, eleven male, and they ranged widely in experience from novice teachers to teachers having over twenty years of classroom experience; 40.4% of the teachers had taught for ten or more years. Teachers that took the survey were from more than twenty-three different schools drawn from more than sixteen different school districts distributed across the state. Although an initial attempt at collecting a large random sample was made, and indeed the final sample bears similar demographic characteristics to that of the teacher population of the state, the final sample of 42 teachers was altogether one of convenience as all contacts were made with teachers through the ultimate mercy of cooperative school administrators. Brief results regarding the reliability and validity of the measures obtained follows.

Reliability

As explained above, the PCK-MELL survey was built from a framework hypothesizing three underlying factors: OBST, RESRC and STRAT. Hence, the computation of Cronbach's (1951) α , a standard measure of internal consistency for surveys and which assumes the unidimensionality of the instrument, using all of the items as a single factor would in fact yield an inappropriate measure of reliability. Nevertheless, as this survey was exploratory in the sense that both the construct, the framework and the items were all novel, this alpha was computed at the outset as a starting point for further investigation of internal consistency. Cronbach's alpha, equivalent to Kuder-Richardson's formula 20 for dichotomously scored items (correct or incorrect), was .431 for the whole set of 32 items of the instrument. This alpha was taken as the first potential confirmation of multidimensionality.

To better understand the factor structure of the test, a combination of evidences from the theoretical orientation of the items, along with from Item Response Theory (IRT), factor analysis, and Cronbach's alpha were used to separate the items into different scales. First, the following two-parameter IRT model was used to compute difficulty and discrimination

coefficients for the full set of 32 items: $p(x_j = 1 | \theta, a_j, \delta_j) = \frac{e^{a_j(\theta - \delta_j)}}{1 + e^{a_j(\theta - \delta_j)}}$. In comparison with

classical test theoretic (CTT) measures, like Cronbach's alpha, which takes the observed score on the entire instrument as the unit of analysis, IRT models take the item as the unit of analysis (De Ayala, 2009). IRT is an important tool in psychometric test development. However, like Cronbach's alpha, IRT models assume unidimensionality. Hence, the initial IRT model was unstable, with about an equal number of items having positive discrimination coefficients and negative discrimination coefficients. The sign of these coefficients, positive or negative, was

then used to relegate items to one of two scales, the first concerned more with the knowledge of both obstacles and resources (OBST/RESRC) and the second concerned more with knowledge of strategies (STRAT). The numerical assignment of items (using the positive or negative sign of the discrimination coefficient) held a high degree of agreement with the theoretical orientation of the items. That is, items with positive coefficients were more closely aligned theoretically with each other than with items having negative coefficients. Furthermore, the dual classification of items achieved through the IRT coefficients was also verified by confirmatory factor analysis (Principal Component Analysis with Promax rotation) in the case of all but one of the items. Finally, eleven (11) items were excluded from this analysis entirely because of weak correlation (Pearson r) with items in either of the two scales. The resulting two scales, OBST/RESRC and STRAT, had Cronbach's alpha reliabilities of .621 and .606, respectively. These two scales were used in the following assessment of the validity of the measures.

Validity

Measurement validity of the scales was investigated using the concept of a nomological net (Cronbach & Meehl, 1955) and its implied measures of convergent (and discriminant) validity (Campbell & Fiske, 1959). The survey included several questions related to teachers' education and licensure, to teachers' experience, and to teachers' linguistic knowledge, variables hypothesized in this study to be related with knowledge for teaching ELLs. The teachers' scores on the two scales (OBST/RESRC and STRAT), computed as a simple percentage of "correct" responses, were then regressed linearly on these variables. After entering all of the teacher variables, the percentages of variance in scores on the OBST/RESRC and STRAT scales explained (R-squared) by the models given in the table below were, respectively, 38.6% and 29.5%.

Several observations can be made from Table 2. At a glance it appeared that the OBST/RESRC scale was the more 'difficult' of the two, based upon the linear regression intercept; that is, controlling for all other factors, the average score on the OBST/RESRC scale was 28.7% correct compared to 64% on the STRAT scales. IRT item characteristic and test information curves confirmed this difference in difficulty. Furthermore, Table 2 shows that, for this group of teachers, some of the education and experience variables were significant predictors of differences in scores in the two knowledge scales of the PCK-MELL survey. Teachers that had experience in teaching math classes in which more than 40% of students were ELLs scored 19.4 percentage points higher ($p < .05$) on average on the OBST/RESRC scale than did teachers without this experience. Conversely, teachers that possessed a certification to teach English as a Second Language (ESL) scored 17.1% lower on average ($p < .10$) on this scale than did teachers not having this certification. Teachers that possessed additional certifications to work in other educational settings (e.g., administration) or content areas besides mathematics scored 12.7% higher on average ($p < .10$) on the STRAT scale than did teachers who held mathematics teaching certificates alone. Likewise, teachers that had been teaching mathematics for more than 5 years scored an average of 13.9 percentage points higher ($p < .10$) on the STRAT scale than did less experienced teachers. Finally, teachers that had completed any sort of professional development concerned specifically with ELLs scored 13.7% lower on average ($p < .10$) on the STRAT scale than did teachers not having this experience.

Table 2. *Linear Regression Model of OBST/RESRC and STRAT Scale Scores on Teacher Variables.*

Measure	OBST/RESRC		STRAT	
	Coefficient	SE	Coefficient	SE
Intercept	.287*	.106	.640**	.106
Teacher Education/Licensure				
Have a Degree in Math	-.116	.080	-.054	.080
Certified to Teach Math	.053	.088	.015	.089
Certified to Teach ESL	-.171 [†]	.098	-.008	.099
Had any ELL Professional Development	.013	.067	-.137 [†]	.067
Possess Other Educational Certifications	.074	.076	.127 [†]	.076
Teacher Experience				
Taught Math for More than 5 Years	.066	.072	.139 [†]	.072
Taught More than 3 Different Types of Courses	.122	.095	.071	.096
Taught a Class with More than 40% ELLs	.194*	.072	-.004	.072
Teacher Linguistic Knowledge				
Know Spanish “Well”	-.002	.063	-.004	.063
Speak a Language Other than English or Spanish	.084	.079	-.084	.079
[†] $p < .10$; * $p < .05$; ** $p < .01$				

The PCK-MELL Survey as a Measurement Instrument

PCK-MELL, the underlying hypothetical construct at the heart this survey, was proposed to be a subset of MKT (Ball, Thames, & Phelps, 2008). Furthermore, the construct was initially theorized as composite of three subdomains: knowledge of obstacles that ELLs encounter in math classes (OBST), knowledge of resources that ELLs draw upon to help them learn mathematics (RESRC), and knowledge of instructional strategies for teaching ELLs mathematics (STRAT). Based upon the limited survey data obtained in this study and upon the theoretical relationship between items in the OBST and RESRC domains, the factor structure of the survey seemed to be binary, having not three scales, but two: OBST/RESRC and STRAT.

As was seen above, the reliabilities of these two scales were low by measurement standards. Although some have argued that, for dichotomously scored items, KR-20 greater than .5 is even acceptable (McGahee, T. W., & Ball, J., 2009), alpha of .70 is often considered a lower bound on acceptable internal consistency (Nunnally & Bernstein, 1994) and considerably higher consistency would be needed for commercial usage of such a measure. This means the measurements, and interpretations thereto, obtained in this study must be taken with a degree of skepticism. Nevertheless, many of the items designed to measure knowledge of content and students (KCS), one of the larger domains to which the OBST/RESRC and STRAT domains pertain, also showed very low reliability during testing, $\alpha < .70$ (Hill, Ball, & Schilling, 2008).

Hence, this study also echoes the complexity experienced by those researchers. Precise measurement of teachers' MKT is indeed a delicate, and at times elusive, accomplishment.

A goal of this paper has been to argue that it may be valid to define knowledge for teaching mathematics to ELLs as a construct subsumed with MKT, and as a part of knowledge for equitable mathematics teaching as a subset of MKT. Although the instrument in its current form was exploratory and is not ready for commercial usage, it may offer a valuable starting point for further theory and instrument development along these lines. Furthermore, some specific directions for improvement are evident. For example, as a relatively inexpensive improvement on reliability, the Spearman-Brown prophecy formula indicates that the mere addition of fewer than ten similar items to each of the scales would elevate reliability to greater than .70 (Brown, 1910; Spearman, 1910). Writing more items is an obvious improvement. But this study also points to the need for more focused research into the specific types of, dimensions of, and relationships between both the *obstacles* and *resources* encountered by ELLs in mathematics. That is, the theory behind this strand of knowledge needs development. Qualitative studies, including careful interviews and observations of teachers in classrooms of ELLs, would seem to offer promising methods to begin to better understand the relationship between ELLs' deficits and affordances, and teachers' understanding of these.

Additionally, the linear regression models presented in Table 2 may offer limited, initial insight into how knowledge in these domains operates. For the OBST/RESRC scale, of the variables potentially related to knowledge in this domain only experience teaching larger numbers of ELLs (greater than 40%) was a significant ($p < .05$) predictor of gain in the scale score for these teachers. Yet, this variable is probably the most important one if this type of PCK is related to actual teaching experience. Similarly for the STRAT scale, experience teaching math for more than five years was a moderate ($p < .10$) predictor of gain in scale score as was possession of multiple teaching certifications. These may, respectively, be proxies for length of teaching experience and breadth of teaching experience. In this case again knowledge in the STRAT domain may be related to actual teaching experience. Finally, it is interesting to note that two of the variables, one in each of the scales, were moderate ($p < .10$) predictors of a *decrease* in scales scores. For the OBST/RESRC and STRAT scales these were, respectively, certification to teach English as a second language (ESL) and whether or not the teacher had experienced any professional development for teaching ELLs. This particular finding is alarming; one would think that both of these variables would relate to increases in knowledge in the domains being tested, not decreases. Yet, if these two variables are seen as proxies for formal *preparation* to teach ELLs, then again the knowledge called upon by this exploratory instrument may be more a matter of actual experience than of formal education. At the least, it would seem that the topics presented in the instrument of this study were different than topics that the teachers in this study had seen in their ELL professional development experiences or ESL classes.

Concluding Discussion

The PCK-MELL survey developed in this study gives an example of a way in which an aspect of knowledge for equitable mathematics teaching has been operationalized in the form of survey items yielding quantitative data and psychometric results that can lead to further theory development. Perhaps most importantly, this study has offered a way of thinking about knowledge needed by mathematics teachers of English Language Learners—and by extension, all second language learners in the mathematics classroom, including immigrant students in other countries—that builds upon and extends current research. Under the framework proposed herein,

in addition to mathematics content knowledge, teachers of ELLs need: knowledge of *obstacles* to learning that are frequently encountered by ELLs in mathematics classes, knowledge of the *resources* that ELLs bring with them to the learning of mathematics, and knowledge of instructional *strategies* for teaching ELLs. This theoretical framework takes the bold, and admittedly controversial stance, of allowing both deficit and affordance perspectives of ELLs as mathematics learners, not advancing one perspective over the other or placing them in opposition, but rather positioning these two perspectives as two *sides of the same coin*: that is, both perspectives may together offer valuable insight for mathematics teachers of ELLs who should be aware both of obstacles to learning that their students are likely to face and also of the resources within and around those very students that can effectively capacitate them to do mathematics and to communicate mathematically. While this way of thinking about learners no doubt applies to all students—regardless of their language background—the viewpoint may be of special value to the particular linguistic experiences that ELLs and second language learners in all language contexts bring to learning mathematics.

This has been an initial attempt at advancing both the theory and the tools that can lead to better understanding teacher knowledge factors that impact achievement in mathematics among ELLs. The article offers an example of an exploratory process that was used to lay the bearings of a theoretical framework of knowledge needed for teaching mathematics to ELLs and also of some steps that were made toward developing a research tool to be used to study this knowledge. This study indicates the inadequacy of current theory and tools for studying aspects of mathematics teachers' knowledge related to equitable access for diverse student populations. It is hoped that this paper may make a small contribution toward bringing more clearly into focus a research agenda from which a viable theory of knowledge for equitable mathematics teaching in the context of linguistically and culturally rich and diverse populations of mathematics students will come forth.

References

- Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, 59(5), 389–407.
- Brown, W. (1910). Some experimental results in the correlation of mental abilities. *British Journal of Psychology*, 3, 296–322.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56(2), 81–105.
- Capps, R., Fix, M., Murray, J., Ost, J., Passel, J., & Herwantoro, S. (2005). *The new demography of America's schools: Immigration and the No Child Left Behind Act*. Washington, DC: The Urban Institute.
- Chang, M. (2008). Teacher instructional practices and language minority students: A longitudinal model. *Journal of Educational Research*, 102(2), 83–97.
- Coggins, D., Kravin, D., Coates, G. D., & Carroll, M. D. (2007). *English language learners in the mathematics classroom*. Thousand Oaks, CA: Corwin Press.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297–334.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281–302.
- Cuevas, G. J. (1984). Mathematics learning in English as a second language. *Journal for Research in Mathematics Education*, 15(2, Minorities and Mathematics), 134–144.
- De Ayala, R.J. (2009). *Theory and practice of item response theory*. New York, NY: The Guilford Press.
- DeVellis, R. F. (2003). *Scale development: Theory and applications*. *Applied social research methods series*, vol. 26. Thousand Oaks, CA: Sage.
- Dillman, D. A., Smyth, J. D., & Christian, L. M. (2009). *Internet, mail, and mixed-mode surveys: The tailored design method*. Hoboken, NJ : John Wiley & Sons, Inc.
- Echevarría, L., Short, D., & Vogt, M. (2007). *Making content comprehensible for English learners: The SIOP model* (3rd ed.). Boston: Pearson Educational Press.
- Francis, D., Rivera, M., Lesaux, N., Kieffer, M., & Rivera, H. (2006). *Practical guidelines for the education of English language learners: research-based recommendations for instruction and academic interventions*. Portsmouth, NH: RMC Research Corporation, Center on Instruction.
- Gutiérrez, R. (2002). Beyond essentialism: the complexity of language in teaching mathematics to latina/o students. *American Educational Research Journal*, 39(4), 1047–1088.
- Gutiérrez, R. (2008). A "gap-gazing" fetish in mathematics education? Problematizing research on the achievement gap. *Journal for Research in Mathematics Education*, 39(4), 357–364.

- Gutstein, E., Lipman, P., Hernandez, P., & Reyes, R. d. I. (1997). Culturally relevant mathematics teaching in a Mexican-American context. *Journal for Research in Mathematics Education*, 28(6), 709–737.
- Henderson, R. W., & Landesman, E. M. (1995). Effects of thematically integrated mathematics instruction on students of Mexican descent. *The Journal of Educational Research*, 88(5), 290–300.
- 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., Ball, D. L., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education*, 39(4), 372–400.
- Hill, H.C., Ball, D.L., Sleep, L. & Lewis, J.M. (2007). Assessing teachers' mathematical knowledge: What knowledge matters and what evidence counts? In F. Lester (Ed.), *Handbook for Research on Mathematics Education (2nd ed)* (pp. 111–155). Charlotte, NC: Information Age Publishing.
- Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D.L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction*, 26(4), 430–511.
- Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, 42(2), 371–406.
- Hill, H. C., Schilling, S. G., & Ball, D. L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *Elementary School Journal*, 105(1), 11–30.
- Khisty, L. L., & Morales Jr., H. (2004). Discourse matters: Equity, access and Latinos learning mathematics. Retrieved September 1, 2011 from <http://www.icme-organisers.dk/tsg25/subgroups/khisty.doc>
- Lager, C. A. (2006). Types of mathematics-language reading interactions that unnecessarily hinder algebra learning and assessment. *Reading Psychology*, 27(2–3), 165–204.
- Llabre, M. M., & Cuevas, G. (1983). The effects of test language and mathematical skills assessed on the scores of bilingual Hispanic students. *Journal for Research in Mathematics Education*, 14(5), 318–324.
- Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Mahway, NJ: Erlbaum.
- Martiniello, M. (2009). Linguistic complexity, schematic representations, and differential item functioning for English language learners in math tests. *Educational Assessment*, 14(3–4), 160–179.
- McGahee, T. W., & Ball, J. (2009). How to read and really use an item analysis. *Nurse Educator*, 34(4), 166–171.

- Meskill, C. (2005). Infusing English language learner issues throughout professional educator curricula: The training all teachers project. *Teachers College Record*, 107(4), 739–756.
- Moschkovich, J. N. (2002). A situated and sociocultural perspective on bilingual mathematics learners. *Mathematical Thinking and Learning*, 4(2 & 3), 189–212.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.
- Payán, R. M., & Nettles, M. T. (2008, January). *Current state of English-language learners in the U.S. K–12 student population* (English-Language Learners Symposium Fact Sheet). Princeton, NJ: Educational Testing Service.
- Schon, D. A. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14.
- Schuman, H., & Presser, S. (1996). *Questions and answers in attitude surveys: Experiments on question form, working, and context*. London: Sage.
- Sorto, M. A., Mejía Colindres, C. A., & Wilson, A. T. (Sept., 2014). Uncovering and Eliciting Mathematical Perceptions in Linguistically Diverse Classrooms. *Mathematics Teaching in the Middle School*. 20(2), pp. 72-77.
- Spearman, Charles C. (1910). Correlation calculated from faulty data. *British Journal of Psychology*, 3, 271–295.
- U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, (2012). *The condition of education: Participation in education – elementary/secondary enrollment*. (Table-A-6-1. Number and percentage of children ages 5-17 who spoke only English at home, who spoke a language other than English at home and who spoke English with difficulty, and percent enrolled in school: Selected years, 1980 – 2009). Retrieved from <http://nces.ed.gov/programs/coe/tables/table-lsm-1.asp>.
- Watson, S., Miller, T. L., Driver, J., Rutledge, V., & McAllister, D. (2005). English language learner representation in teacher education textbooks: A null curriculum? *Education*, 126(1), 148–157.
- Wolf, M. K., Herman, J. L., Dietel, R. (2010). *Improving the validity of English language learner assessment systems. Policy brief 10, spring 2010*. National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

In-service Teachers' Reasoning about Scenarios of Teaching Mathematics to English Language Learners

Sultan Turkan
Educational Testing Service

Abstract: The student population in the U.S. and worldwide is becoming increasingly diverse, creating a need to support *all* learners, especially linguistically and culturally diverse subpopulations such as English language learners (ELLs). From a social equity standpoint, the need to support these learners is critical especially in mathematics classrooms. In the U.S, the demand for mathematics teachers who are adequately prepared to teach ELLs has in fact risen. Yet, little is known about what knowledge base is essential to teach mathematics to ELLs. Driven by the need to explore this knowledge base, in this paper I explore what is involved in reasoning about teaching mathematics to ELLs. To this end, a set of instructional scenarios illustrating the work of teaching mathematics to ELLs was utilized within an assessment environment. Interviews with 10 mathematics teachers reasoning about the scenarios showed that they drew on the information provided about ELLs' proficiency levels while reasoning through the scenarios. Also, teachers' reasoning seems to be qualified by the extent to which they could both use their content knowledge in mathematics and modify their instructional choices according to ELLs' language needs specified in the scenarios. This study motivates large-scale future studies examining what systematic teacher knowledge base might differentiate good teaching for ELLs from good teaching for *all* students.

Key Terms: Teaching mathematics, English language learners, teacher knowledge base, and instructional scenarios.

Introduction

As the student population in U.S. classrooms becomes increasingly diverse, it is critical that all teachers develop the knowledge and skills to support English language learners (ELLs) in mainstream classrooms at every grade level. However, the majority of teachers are not adequately prepared to teach academic content to ELLs (Lucas & Villegas, 2011). Mathematics may be especially challenging for ELLs (Martiniello, 2008). The National Assessment of Education Progress (NAEP) mathematics test results in 2013 (United States Department of Education, 2015) showed that the achievement gap between non-ELL and ELL students both at the 4th and 8th grades did not significantly differ from the achievement gaps reported in previous years in 2011, 2009, 2000, or 1996. Challenges in learning mathematics might preclude access to mathematical and scientific fields, which raises a critical issue of equity (Moschkovich, 2002; Oakes, 1990; Secada, 1992). Teachers play a key role in leveling the playing field for ELLs in mathematics (Moschkovich, 1999; Secada, 1998). However, there is a clear gap between the demands of teaching ELLs and the supply of mainstream mathematics teachers who are

adequately prepared to teach (Ballantyne, Sanderman, & Levy, 2008). As the ELL population has increased up to 10% of public school students (National Clearinghouse for English Language Acquisition, 2011), the demand for trained mathematics teachers has as well; many, if not most, teachers need to develop the knowledge and skills needed to teach in classrooms with increasingly diverse populations of students.

Despite the need, little is known about the knowledge base needed to teach mathematics to ELLs. Recent conceptualizations of the teacher knowledge base (Ball, Thames, & Phelps, 2008) have not addressed how manifestations of it change depending on who the students in the classroom are and what they bring to the learning experience. To inform research on the teacher knowledge base, most of the existing scholarship on educational linguistics (de Oliveira & Cheng, 2011; Fang, 2006; Lucas, Villegas, & Freedson-Gonzalez, 2008; Schlepppegrell, 2001) suggests that language is central to teaching mathematics and that mathematics teachers, even at the secondary level, should know how to identify the linguistic demands in the mathematical content and model mathematics-specific language use. This line of work suggests that language serves as a medium to understanding and communicating the content in mathematics, which in turn is central to the work of teaching mathematics to ELLs.

However, our understandings as to what this knowledge base entails and how it gets or should get enacted in mathematics classrooms are limited. To better understand the knowledge base and reasoning involved in teaching mathematics to ELLs, eliciting teachers' reasoning around particular scenarios formatted in an assessment environment might help. As research (Ball & Hill, 2008; Jacobson, Remillard, Hoover, & Aaron, in press) suggests, there needs to be a reciprocal interplay between conceptualizations of teacher knowledge base and measurement in that the development of teacher knowledge assessments needs to be informed by conceptual work in the domain of effective and equitable teaching of mathematics. To facilitate this interplay, new lines of research also show that authentic classroom scenarios formatted into the assessment environment could serve as powerful tools for eliciting teacher reasoning as well as reflective teacher learning (Lai & Howell, 2014). Further, emerging research shows that instructional scenarios help to situate explorations about pre-service and in-service teachers' reasoning about teaching (Grossman et al., 2009; Masingila & Doerr, 2002). This study was therefore based on the premise that the instructional scenarios might be instrumental in explorations of the knowledge base needed to teach mathematics to ELLs.

Driven by the need to explore essential knowledge for teaching mathematics to ELLs, in this study, we utilized a set of instructional scenarios embedded in an assessment environment to explore what is involved in reasoning about teaching middle school mathematics to ELLs. First, I developed a theoretical lens explicating how understanding the language demands in any content area is central to teaching the content in diverse ways to diverse learners. Second, we developed scenarios illustrating the work of teaching mathematics to ELLs in an assessment environment and conducted interviews with mathematics teachers to explore their reasoning about the scenarios. The current paper is part of a larger study (Turkan, Croft, Bicknell, & Barnes, 2012) and sought to address the following research questions. In addressing these questions, teachers' strong and weak reasoning about the given instructional scenarios as evident in cognitive interviews was contrasted as a way of anchoring the exploration of knowledge essential for teaching mathematics to ELLs.

1. How do I conceptualize the knowledge for teaching mathematics to ELLs?
2. How do practicing mathematics teachers reason about the teaching of mathematics to ELLs?

Towards addressing the first research question, the next section provides an overview of the theoretical framework. The framework informed the development of the measure. Description of the measure and methods is followed by the discussion of results addressing the second question about how a sample of mathematics teachers reasoned through the teaching of mathematics to ELLs. Finally, directions for future research are shared.

Theoretical Lens

Our assumption in identifying and proposing a knowledge base was that all teachers need to develop a specific knowledge base in order to make academic content accessible and meaningful to ELLs. This assumption has long been an argument among scholars that all teachers need to be provided with opportunities to develop an adequate knowledge base to make academic content accessible and meaningful to ELLs (de Jong & Harper, 2005, 2011; Lucas & Villegas, 2011).

Specifically, effective teachers of ELLs draw upon a specialized knowledge base in order to teach academic content in accessible ways called *disciplinary linguistic knowledge (DLK)* for teaching academic content to ELLs (Turkan, de Oliveira, Lee, & Phelps, 2014). These authors argued that DLK is needed to model for ELLs how language is used to communicate meaning and to engage them in disciplinary discourse. That is, DLK is a specialized knowledge that teachers need to have regarding how language is used to communicate the ideas and concepts within a particular disciplinary discourse. Discourse, in this context, refers to ways of knowing, constructing, and communicating knowledge. Involved within the discourse of an academic discipline is the academic register of the discipline that includes linguistic features specific to that discipline. DLK is an elaborated version of Reeves' (2009) use of the term, *Linguistic Knowledge for Teaching*, which is operationalized as the linguistic knowledge that English for Speakers of Other Languages (ESOL) teachers use to create opportunities for learners to communicate meaning. In this paper, however, I use DLK to refer to knowledge that *content* teachers need to facilitate ELLs' content understanding.

Additionally, this knowledge base is a facet of, but not equivalent to, content knowledge for teaching, CKT, (Ball et al., 2008). While the body of work on teacher knowledge measures and research on content knowledge for teaching (Ball, Hill, & Bass, 2005; Ball et al., 2008) is growing, it does not address the role of special student populations in how teachers attend to the tasks of teaching in the classroom. More specifically, CKT doesn't address what knowledge base content teachers draw on to teach special student populations such as ELLs. Outside general teacher education literature, several scholars (Bunch, 2013; Galguera, 2011) have conceptualized what knowledge base teachers need to teach content to ELLs. Bunch accounts for the pedagogical aspect of teaching content to ELLs while by proposing DLK, Turkan et al. (2014) complement Bunch's view that content teachers should also address ELLs' linguistic needs at the word, sentence, and discourse levels as these levels pertain to a particular discipline.

DLK encompasses teachers' knowledge of content-specific academic discourse as well as their ability to present conceptual language in multiple ways that are accessible to ELLs. Specific

tasks of teaching subsumed therein include two subdomains: 1) identifying linguistic features of the disciplinary discourse and 2) modeling for ELLs how to communicate meaning in the discipline and engaging them in using the language of the discipline orally or in writing. These tasks are conceptualized further below. The measure of teaching quality for ELLs was developed (Turkan et al. 2012) according to this conceptualization of the targeted construct.

Identifying Linguistic Features

The argument for the first subdomain is that DLK entails particular knowledge for identifying and unpacking the linguistic features and language demands of a disciplinary discourse to make the content accessible to ELLs. We argue that, as part of the unique knowledge base for teaching ELLs, teachers know which linguistic features associated with the particular content might cause or constitute ELLs' misunderstandings or misconceptions about the content. That is, to resolve misconceptions and facilitate ELLs' comprehension and interpretation of the content, linguistic features need to be unpacked. There are two premises in this argument: 1) there are linguistic features that convey meaning and content in a given discipline, 2) the linguistic choices are made in each discipline at the word, sentence, and discourse level to convey meaning.

The first premise is that teachers identify the linguistic features that are specific to a content area. In doing so, they decode or unpack the linguistic features of the discipline to build connections between content and meaning and particular linguistic features and structures that convey the particular meanings. By 'unpack', we mean that teachers make the linguistic *form-meaning* connections of the disciplinary discourse explicit for students. In other words, teachers make the content-related input comprehensible and accessible to ELLs (Cummins, 2000). *Form* here refers to a string of words grammatically put together to carry *meaning*. A string of forms (words, sentences) might carry various meanings depending on the context in which it is being used. For instance, McCarthy (1991) discusses how the following sentences might carry the meaning of question, statement, or command: "You don't love me: (a) question (b) statement" or "you eat it: (a) statement (b) command" (p. 9). Form-meaning connections (Schleppegrell, 2013) are made when teachers attend to the features of language (i.e., form) and they simultaneously model for ELLs the ways in which meaning and content is communicated in the particular discipline. For example, linguistic expression such as "three times as big as..." has a particular correspondence in mathematics, hence a certain mathematical form-meaning connection. As teachers identify the linguistic choices made to convey meaning in a particular discipline like mathematics, they also make the content comprehensible to ELLs by explicitly teaching specific language functions, forms, or meaning behind the text in order for ELLs to learn how the linguistic choices are being used to convey the particular meaning. Once ELLs can differentiate and be made aware of the linguistic choices through the comprehensible input received, they can more readily participate in producing or using the language orally or in writing to convey their understandings about the content.

The second premise is that the ways in which meaning is communicated in content areas are instantiated through linguistic choices, for example, at vocabulary, grammatical and syntactical structures, and discourse levels (Christie & Martin, 1997; Gebhard, Willett, Pablo, Caicedo, & Peidra, 2011; Schleppegrell, 2004). The linguistic choices operate at the word, sentence, and discourse levels: at the word (e.g., *square root* of 25), sentence (e.g., *taking the square root is the inverse operation of squaring*), and discourse (e.g., *taking the square root*

involves finding the number that, when multiplied by itself, gives 25) levels. The linguistic choices at these three levels exist to communicate meaning and perform language tasks and functions associated with the particular discipline such as writing laboratory reports in the science classroom (Schleppegrell, 2004), explaining solution processes or describing conjectures in the mathematics classroom (Moschkovich, 1999), writing personal recounts of an event in the English language arts classroom (Brisk & Zisselberger, 2011), and retelling events or presenting debates in the social studies classroom (Fang & Schleppegrell, 2008).

The above premises led us to argue that teachers should be able to identify the linguistic features specific to a content area so that they can decode or unpack the linguistic features of the discipline and build connections between content and meaning, on the one hand, and particular linguistic features and structures that convey the particular meanings, on the other; and that teachers should be able to explicitly teach what constitutes appropriate linguistic choices in their discipline.

Modeling How to Communicate Meaning Orally or in Writing

The argument for the second subdomain is that DLK involves teachers' knowledge for modeling the ways in which the discourse of a discipline is constructed and for engaging ELLs in communicating meaning in the disciplinary discourse orally or in writing. When modeling the disciplinary discourse, teachers build on the mastery of the first domain to make the linguistic features of a content area explicit for ELLs. Further, teachers draw on DLK to engage ELLs in learning how the rules of the linguistic features function to convey meaning in the content area. In doing so, teachers encourage ELLs to explore and build *form-meaning* connections to read, write, listen, speak, and think in the language of the discipline. In this process, students produce work in writing or orally in which they demonstrate their knowledge of the discipline using the disciplinary language as it was modeled to them. Hence, students participate in using the language of the content for complex academic tasks such as "to generate new knowledge, create literature and art, and act on social realities" (Cummins, 2000). In the process of constructing *form-meaning* connections, students develop awareness of and knowledge about what linguistic features are used to represent the meanings and ideas associated with the content.

In this section, we have proposed that DLK encompasses teachers' knowledge of disciplinary discourse in order to represent the disciplinary content in accessible ways to ELLs. I have also argued that to make content accessible, every teacher needs to be able to identify and unpack the ways in which the linguistic features are connected to the meanings of a discipline. As teachers raise ELLs' awareness explicitly around how form-meaning connections are linked in the discourse of the discipline, they will be able to engage ELLs in reading, writing, listening, speaking, and thinking in the disciplinary discourse. Therefore, within the purview of DLK, teachers should know about the linguistic choices made in a given discipline and should be able to identify the language demands and unpack the meaning-form relationships for ELLs to model for them and engage them in using the discourse of the discipline.

Methods

To explore and identify the knowledge base needed to teach mathematics to ELLs, I have grounded my study in practice-based reasoning about teaching. Specifically, teachers' reasoning about the given scenarios of teaching mathematics to ELLs was elicited through interviews. In

doing so, I also explored how the high and low performing teachers, as identified by their number correct scores, reasoned about the scenarios. The scenarios were designed as part of the operationalization of the theoretical framework described above. Scenarios were formatted as part of the assessment development process, which served as a means to operationalizing and refining the domain. Next, the assessment development process is explained; later, the interview data collection procedures are presented along with the analysis.

The assessment development process started with the identification and verification of the domain of teacher knowledge and skills required to teach ELLs mathematics content effectively. I initially identified this domain through a review of the academic literature base on teaching mathematics to ELLs, as well as review of the state standards for teacher certification in the three content areas: mathematics, science, and English language arts. Based on the general and content-specific literature, we developed a set of 67 statements describing what content teachers should know and be able to do (for more information, see Turkan et al., 2012). Based on the description of the targeted domain, we categorized these statements into two subdomains: pedagogical knowledge and linguistic knowledge. Under each domain, there were statements generic to all content areas and specific to particular content areas. To validate these statements, we conducted a national survey of practitioners and teacher educators, receiving 269 responses. Subsequently, a panel of 14 teacher educators and teachers further validated the statements by reaching consensus that the statements supported the claims of the assessment under development (for more information about the panel, see Turkan et al., 2012). The principles of evidence-centered design (ECD; Mislevy, 1994; Mislevy, Almond, & Lukas, 2003) guided the panelists' review of the statements about linguistic and content knowledge necessary to teach ELLs. The panelists helped to unpack specifically what teachers would need to perform to manifest that they have the essential knowledge and skills. Based on the performance indicators the panel helped to further specify, the assessment was developed to measure the validated domain of teacher knowledge essential to teach content to ELLs.

During assessment development, we identified the language demands inherent to the content standards and topics selected from the 6th and 8th grade mathematics curricula. We selected middle grades based on the content and language demands—and informed by the panelists-- we identified a set of instructional practices representative of effective ELL teaching. With that, we developed authentic instructional scenarios embedded in an assessment environment (see appendix for two sample items). The scenarios included variant and nonvariant features. Nonvariant features were the effective mathematics teaching practices targeted for assessment. Variant features included description of the learning objective, the ELL student's task, as well as the characteristics of the focus ELL(s), which include their ELL proficiency level descriptors. Various indicators of ELL proficiency in at least one of the four skills (reading, writing, speaking, and listening) were specified. These indicators were provided to guide teachers' understanding about what ELL(s) of particular focus could do or could not do. The learning or teaching objectives were informed by the targeted 6th and 8th grade mathematics standards. This then determined the language demands that are embedded in the particular learning or teaching objectives. The language demands guided what tasks ELLs, characterized by specific linguistic skills in the instructional scenario, were confronted with in the classroom. For example, item 1, illustrated in the appendix, focused on the mathematical concept of 'distributive property' in a classroom scenario in which teachers are faced with a choice to either use the definition that they find or to provide an illustration of the concept to make it accessible for the ELLs described in the scenario. This focus then determined the language demands in explaining

and representing the concept for understanding by the specified ELLs. Similarly, item 2, discussed in the results section, focused on a learning objective; namely, writing and solving equations based on word problems. In the particular classroom scenario, ELLs were characterized as having grade-level content knowledge but needing linguistic support in understanding and solving the given word problem. This item is representative of the other items developed to assess teachers' knowledge of content for identifying the language demands embedded in the content. That is, most of the items focused on teachers' knowledge and skills for identifying and connecting the content and language aspects of the content knowledge for teaching.

Approximately 60 items were developed to assess the DLK needed to teach mathematics to ELLs. Twenty of the 60 mathematics items were selected as a representative sample of the domains targeted for assessment. The sample was administered online to 60 middle school mathematics educators who teach in school districts that serve large numbers of students who are ELLs. Samples of mathematics teachers represented southern, western, and northeastern regions of the United States. These teachers had a minimum of five years teaching experience in predominantly ELL school districts. The initial analysis revealed a Cronbach's alpha ($\alpha = .50$) and only five items showed biserial correlation measures of item discrimination below 0.2.

Interviews

To address how teachers reason about ELL teaching, we invited all 60 in-service mathematics teachers who took the assessment for a follow-up interview. Only 27 indicated interest to participate in the interview. We ranked the test performance of these 27 potential interview participants according to their number-correct scores. Two mathematics teachers got 14 while seven teachers got 11 of the 20 items correct, which was the mode on the test. Twelve teachers got fewer than seven items correct. Low and high performing teachers were identified by the number of items they got right out of the 20 multiple choice items. Those who got 10 and above items right were considered high performing while other teachers who got 10 and below items right were considered low performing on this particular test. This cut off decision was arbitrary based on the distribution of the number-correct scores. Based on this decision, we randomly selected five high performing and 5 low performing teachers, who agreed to participate in the interview. All 10 teachers had taught mathematics for a minimum of five years at predominantly ELL school districts where each teacher had at least five ELLs in their respective classrooms.

The goal of the semi-structured phone interviews was to allow teachers to explain their reasoning as each of them recalled their response to a sample of six items randomly selected from the 20 items. The aforementioned five items with low item discrimination characteristics were not included in the interviews. The remaining items were distributed across the interview participants in a way that each eligible item would be discussed by two teachers. The selected group of items was sent to the participants at least 24 hours ahead of the interviews so that they could have the chance to explain their reasoning about the items. The goal of the interviews was to elicit teachers' reasoning about the items through the use of retrospective questions (see appendix) more than to evaluate or elicit their ability to recall their exact reasoning at the time of taking the assessment.

Analysis

I coded the transcriptions of the interviews using NVivo 9 software. The a priori themes were 1) relevance of scenarios to teachers' work within the specified domains (i.e., identifying linguistic features, modeling how to communicate meaning orally and in writing), and 2) relevance of domains to the items. The subsequent analysis followed an "expanding frame" in which the analysis of the qualitative data began with a tight focus on one element or a few elements (Lindlof, 1995). Within this frame, as the researcher collects evidence, and sees new ways to consider, the frame of evidence is widened in analysis. Specifically, the analysis initially focused on understanding the extent to which teachers considered the ELL proficiency levels identified in the given scenario. As I tried to further address what is involved in this, other themes emerged such as teachers' views about ELL teaching and reasoning about the language demands of the mathematical content. I noted the themes as they came up across participating teachers and the items they discussed during the interviews. Comparison of reasoning between high and low-performing teachers helped to further explain the patterns noted across the participating teachers.

Results

In response to the research question, there were three main findings. One is that seven teachers agreed on considering what ELLs can or cannot do in their instructional decisions. Three teachers argued that all learners (i.e., not just ELLs) have difficulty with the language of mathematics. With this assumption, one teacher claimed that good ELL teaching is effective for all learners. The second finding is that the depth of teachers' content knowledge and sensitivity about ELLs' language needs played a noteworthy role in their reasoning. This may have been a consequence of the particular scenario in the assessment that was used for the interview. If the scenario elicited not only teachers' content knowledge but also their sensitivity about ELLs' language needs, the depth of teachers' content knowledge seemed to surface more noticeably. Third, all ten teachers believed that one effective way to teach mathematics to ELLs was to use pictures as a way of simplifying language and/or removing any language demands specific to mathematics.

Role of ELL Proficiency Indicators

Cognitive interviews with mathematics teachers who either performed well or poorly on the test questions revealed how teachers drew on the information about students' proficiency levels or other information given in the instructional scenarios to reason through the scenarios. Both high-and low-performing teachers drew on the information in their reasoning across all the six scenarios that they discussed during the interviews. I recorded the presence or absence of teachers' reference to the information provided in the scenario. When asked to rate the level of importance for identifying ELLs' proficiency levels and explain why, seven teachers agreed consistently across all the scenarios assigned to them that it was highly important to modify instruction accordingly. However, three of the low-performing teachers consistently explained across the six scenarios that language proficiency descriptors were of low relevance to mathematics instruction and they did not consider what ELLs can or cannot do in identifying the most appropriate instructional practices. One of these particular teachers' assumptions was that

language of mathematics was challenging for all students and there is nothing special to the work of teaching mathematics to ELLs.

One teacher did not see the relevance of understanding what students could or could not do in their second language. The teacher reasoned that mathematics itself is a second language for everybody, not just for ELLs, and therefore ELLs should not receive special consideration during instruction above and beyond non-ELLs.

I(interviewer): Okay, on a scale of 1 to 5 with 5 being the most important, how important do you think it is for math teachers of ELLs to identify the language proficiency levels of their students?

R(espondent): I think, well I think that; I would call it about a 2, meaning not as important, if I got your scale right. Because, you know, the way I think of math, and I tell them this, is that math is a language all by itself, and so everybody learning math is in fact kind of learning a second language. So in a math class, they may be using English to learn an additional language, that being math, so everybody's kind of at a disadvantage in a way.

Along similar lines, these three teachers thought that effective ELL teaching is good teaching for everyone. This reasoning might bear two related interpretations. One might be that teachers think that there is no uniqueness to the quality of teaching mathematics to ELLs. Another is that if a particular effective teaching benefits ELLs, it would benefit everyone. One teacher who represents this line of reasoning stated:

R: No, but you know the thing is is that the way I conduct my classes, when I am doing whole group instruction, or we have a word problem, I basically teach those the same way whether I have ELLs in my class or not, in that we read them out loud, and making sure that even some kid that was raised in America on English may not understand some of the idiomatic ways of expressing ourselves that sometimes creep into these questions. So sometimes you know whether I have ELLs or not, I spend a second making sure that everybody understands what the question says. And the pictures and some of those things that help ELLs help everybody. So if you, to me, if you teach with ELLs in mind in general, you're helping everybody.

All in all, teachers found it helpful to consider what ELL can or cannot do in their instructional decisions. A few teachers, though, claimed that language of mathematics is difficult for all learners and that there is no distinctness to the work of teaching ELLs.

Teachers' Reasoning about ELL Teaching

As for teachers' reasoning about the scenarios, there were several dynamics that played a role. One was the content and focus of the items. Since the items focused on teachers' content knowledge and guided teachers to pay attention to ELLs' proficiency levels, the strength of teachers' content knowledge determined the depth of their reasoning. This relationship was observed in varying degrees along a continuum of depth in all 10 teachers' responses. Here, the continuum of depth in teacher reasoning is presented by illustrating a few sample responses from high and low performing teachers.

To illustrate, item 1 (see appendix) contextualized the teacher's task as introducing the distributive property and providing examples to scaffold students' understanding of this property.

The teacher, in this scenario, has to have a good understanding of the concept ‘distributive property’ to make the connection that options A and B are two different but accurate representations of the same concept. One representation utilizes mathematical symbols (option A) while the other explains the concept with numerical and graphic illustrations. Item 1 assessed teachers’ ability to represent the concept in relation to the particular characteristics of the students, tapping into knowledge of content and students. In terms of student characteristics, the scenario specified that the teacher has newly arrived ELLs whose placement test results indicate that they are weak in mathematics and low in English proficiency. Given the ELLs in the class, the teacher in the scenario decides to introduce distributive property with an illustration, instead of starting the lesson with a linguistically complex definition of distributive property. One of the teachers participating in the interview who had performed low on the overall assessment did not choose option B, intended key, because to her, *there was too much going on in that illustration* (see the original line from her below) and so it would be hard for ELLs to follow. According to this teacher, option A, on the other hand, conveyed the central conceptual understanding behind distributive property. However, this teacher seemed to have missed the nuance that option B conveyed the same conceptual understanding except with more illustration, which was more helpful for the particular ELLs than option A, as well as the more linguistically demanding illustrations given in options C and D. Option A was not helpful for the particular needs of the ELLs with weak mathematics and language skills as it just offered a symbolic representation of the definition for distributive property and so did not serve the instructional intentions of the teacher mentioned in the scenario. We observed that by choosing option A, the participating teacher did not combine her content knowledge and linguistic needs of the ELLs in the instructional scenario.

R: I’m almost positive A was my answer on the thing as well.

I: Okay. And why not B?

R: Too much going on. I just think there’s way too much going on there. A is a clear example where you can see how the number outside the parentheses was moved to meet each of the numbers inside the parentheses. I just think showing them all those blocks, there’s way too much going on there.

I: And you think A is the simplest way.

R: Um hmm.

I: But what other information in the item did you draw on when selecting A?

R: When selecting A?

I: Yeah.

R: Because they don’t want to, they want to illustrate the distributive property and the distributive property says that you multiply the number outside the parentheses the number inside and add it together. I just think A clearly states that, you know, showing the mathematical symbols and everything like that and showing how one side of the equal sign equals the other.

Another teacher who scored higher on the test was swift in combining the characteristics of ELLs in the scenario with her content knowledge about mathematical representations of distributive property. This provided her the affordance to be able to reason that low English

proficiency kids would still get lost with the teacher explanation that option A calls for, unless they were strong in mathematics and she eliminates this option.

I: Can you tell us in your own words what information from the question helped you to select an answer?

R: Basically the characteristics of the focus ELLs. In most questions, that was the thing that I cared the most about was the characteristics of the focus ELLs. These were newly-arrived again. Here, unlike the last one, you gave me more information about them – weak in math – which is not a characteristic of ELLs, that's a characteristic of most kids – and have low proficiency in English. So these are low proficiency English kids which meant to me that they needed pictures. You know the instructional scenario, there's the definition, the definitions, mathy definitions are horrifyingly difficult to follow even for people who speak English.

R: And then when I got down to the choices, you know A is okay because, again, math itself is a language and I can, as a math teacher, clearly say that that's essentially the definition of the distributive property. But really I knew I didn't want to go straight into word problems down in C and D and using those pictures and B, which is my answer choice, that that seemed to me clearly the way to go.

Furthermore, seven teachers, regardless of how well they performed on the test, consistently conceived ELLs as the special student population that constantly needs linguistic support in the form of either linguistic simplification or removal of language demands specific to mathematics altogether. I interpreted this conception in a way that language of mathematics and mathematical content are separate entities. When one is present, the other one should be removed. This view, in turn, might have justified all ten teachers' reasoning in favor of using pictures or manipulatives, that is, anything not language-based, in relevant instructional scenarios. In one interview, for instance, the teacher was given a scenario in which a mathematics teacher gives the following word problem to her students (see item 2 in the appendix): "In 1996, the salary of the governor of New York was about \$50,000 less than triple the salary of the governor of Arkansas. The total of the two salaries was \$190,000. Find the 1996 salary of each state's governor." The teacher in the scenario wants to support the ELLs in being able to understand and solve the problem. ELLs in this class have grade level content knowledge, and their English reading proficiency has been identified at WIDA Reading Level 3¹ (there are six WIDA levels). Options include a) providing ELLs with explanations of the difficult phrases in the problem, b) providing ELLs with visual illustration of the problem, c) giving ELLs a version of the problem in simplified English, and d) giving ELLs the equation that the problem is based on and have them solve the equation. The word problem included linguistically demanding phrases such as "\$50,000 less than triple..." and "the salary of the governor of New York" that needed unpacking for the particular ELLs in the scenario to solve the equation. Given the scenario and ELL characteristics, option A was the best instructional strategy because providing a version of the problem in simplified English (option C) or visual illustration of the problem

¹ According to WIDA 6-8 grade cluster can-do descriptors, the ELLs with a WIDA reading level 3 can 1) Identify topic sentences, main ideas, and details in paragraphs; 2) Identify multiple meanings of words in context (e.g., "cell," "table"), 3) Use context clues; 4) Make predictions based on illustrated text; 5) Identify frequently used affixes and root words to make/extract meaning (e.g., "un-," "re-," "-ed"); 6) Differentiate between fact and opinion; 7) Answer questions about explicit information in texts; 8) Use English dictionaries and glossaries

(option B) would not significantly help the ELLs make the meaning-form connection between the mathematical concept and its linguistic representation in the word problem. Along the same lines, it would not be the best strategy to readily give the ELLs the equation (option D) without helping them understand the connection between the word problem and mathematical equation. The low performing teacher we interviewed explained that she would prefer the visuals over all the other options, as stated in the excerpt below. When prompted further, the teacher explains what she understands ‘simplified English’ as referring to “*taking difficult phrases and making them simplified or putting them in simplified forms*” and she says this strategy would be the same as providing explanations of the difficult phrases in the problem. In other words, the teacher discards the two options because she believes them to be identical.

I: So just tell us what information this question gave you and what it asked you to do.

R: Well it gave me a word problem and it first gave me some information about this particular ELL grade level class, that their content knowledge is at grade level and what their reading level is at. And it’s asking which strategy I would use to support them in this word problem.

I: Okay. So what was your answer?

R: I don’t remember. I remember doing it very late at night but I think what I would probably do is a visual and then the next explanation; and I wasn’t sure on this if they want one answer or more than one. Because I might use a combination; now that I’m looking at it, I might use a visual and then explanation because of the difficult phrases.

I: Yeah. If you had to go to choose one, which one would you say?

R: I’d pick visual.

Overall, participating mathematics teachers seem to have taken into account the information about what ELLs can do at particular proficiency levels. Even though the majority believed in scaffolding ELLs’ understanding of mathematics linguistically or otherwise, a few teachers also stated during the interviews that language of mathematics is hard for all learners, not just ELLs. Further, teachers’ content knowledge might play a role in the depth of their reasoning about the language demands embedded in the mathematical content. Another pattern was that effective ELL teaching was equated to removing all the language demands or simplifying language to the extent possible through the use of pictures or visuals.

Discussion

While presence of any systematic teacher knowledge base specific to teaching middle school mathematics to ELLs still requires further investigation, teachers drew on some sources of knowledge to reason through the scenarios. One belief that seemed to exist in teachers’ reasoning is that effective ELL teaching is not unique above and beyond teaching all learners. While this view has conceptually and linguistically been debunked over the years (Harper & de Jong, 2004, 2009; Schleppegrell, 2001), only recently have we come to an empirical understanding that some teachers tend to be more effective with ELLs (Loeb, Soland, & Fox, 2014).

Loeb et al. (2014) examined the extent to which teachers' effectiveness is the same with ELLs as it is with non-ELLs and how much teacher effectiveness changes across classrooms with ELLs and non-ELLs. The premise here is that teacher effectiveness might vary depending on the specific group of students in the classroom. The findings showed that teachers who are effective with ELLs are also effective with non-ELLs and vice versa. However, when the authors regressed student test performance onto teacher characteristics such as Spanish fluency and attainment of a bilingual certification, they found that teachers with a command of Spanish proficiency and bilingual certification were more effective with ELLs, most of whom were Spanish speaking. This finding supports the emerging stance that good teaching for ELLs is characterized differently from good teaching for all (Harper & de Jong, 2004). Based on this understanding, scholars (Bunch, 2013; Galguera, 2011; Turkan et al., 2014) have attempted to identify pedagogical and linguistic aspects of the work of teaching content to ELLs.

From the current small-scale study, it is observed that the majority of the teachers view ELLs as a group of students who need specialized support in learning mathematics. Those who were observed to be skillfully integrating the information given about ELL characteristics with their content knowledge also performed high on the overall assessment. Teachers who had a better understanding about the content and language demands of the material were more flexibly able to reason through the scenarios. All in all, all the participating teachers integrated their existing knowledge or understanding about students, specifically, the given ELL characteristics into their reasoning about the scenarios. However, the extent to which they were able to integrate their knowledge of content and students was limited. One explanation for this might be that the depth of their specialized content knowledge determines the level of flexibility in which they could reason about the mathematical content in relation to its language demands. In other words, a unique contribution of this study might be that there is an interplay between teachers' content knowledge and DLK, which needs further investigation with more mathematics teachers on a larger scale.

These observations about the interviews might suggest that mainstream mathematics teachers need to draw on a systematic knowledge base to facilitate ELLs' linguistic and content-related challenges, as they are not routinely trained to differentiate their instructional practices according to the needs of ELLs in the classrooms. Further research is needed to examine at a larger scale what mathematics teacher knowledge base differentiates good teaching for ELLs from good teaching for *all*.

Conclusion and Future Directions for Research

There is both a growing population of ELLs in U.S. schools and a lack of preparation on behalf of the mathematics teachers who will be teaching them. This paper has barely scratched the surface of identifying and assessing the knowledge and skills needed to teach mathematics to ELLs. Identifying this knowledge base and testing its existence through assessment development is an essential first step towards understanding: What are the most effective pedagogically and linguistically responsive practices for teaching mathematics to ELLs? How can we meet the long term goal of preparing teachers for linguistically and culturally diverse learners? What will constitute quality professional development of mathematics teachers of ELLs?

References

- Ball, D. L., & Hill, H. C. (2008). Measuring teacher quality in practice. In D. H. Gitomer (Ed.), *Measurement issues and assessment for teaching quality* (pp. 80–98). Thousand Oaks, CA: Sage.
- Ball, D. L., Hill, H. H., & Bass, H. (2005). Knowing mathematics for teaching: Who knows mathematics well enough to teach third grade, and how can we decide? *American Educators*, 29(1), 14–46.
- Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, 59(5), 389–407.
- Ballantyne, K. G., Sanderman, A. R., & Levy, J. (2008). *Educating English language learners: Building teacher capacity* (Roundtable Report). Washington, DC: National Clearinghouse for English Language Acquisition & Language Instruction Educational Programs.
- Brisk, M. E., & Zisselberger, M. (2011). We've let them in on the secret: Using SFL theory to improve the teaching of writing to bilingual learners. In T. Lucas (Ed.), *Teacher preparation for linguistically diverse classrooms: A resource for teacher educators* (pp. 55–73). New York, NY: Routledge.
- Bunch, G. C. (2013). Pedagogical language knowledge: Preparing mainstream teachers for English learners in the new standards era. *Review of Research in Education*, 37, 298–341.
- Christie, F., & Martin, J. R. (1997). *Genre and institutions: Social processes in the workplace and school*. London, UK: Cassell.
- Cummins, J. (2000). *Language, power and pedagogy: Bilingual children in the crossfire*. Clevedon, UK: Multilingual Matters.
- de Jong, E., & Harper, C. (2005). Preparing mainstream teachers for English language learners: Is being a good teacher good enough? *Teacher Education Quarterly*, 32(2), 101–124.
- de Jong, E., & Harper, C. (2011). Accommodating diversity: Preservice teachers' views on effective practices for English language learners. In T. Lucas (Ed.), *Teacher preparation for linguistically diverse classrooms: A resource for teacher educators* (pp. 55–73). New York, NY: Routledge.
- de Oliveira, L. C. & Cheng, D. (2011). Language and the multisemiotic nature of mathematics. *The Reading Matrix*, 11(3), 255–268.
- Fang, Z. (2006). The language demands of science reading in middle school. *International Journal of Science Education*, 28(5), 491–520.
- Fang, Z., & Schleppegrell, M. J. (2008). *Reading in secondary content areas: A language-based pedagogy*. Ann Arbor, MI: University of Michigan Press.
- Galguera, T. (2011). Participant structures as professional learning tasks and the development of pedagogical language knowledge among preservice teachers. *Teacher Education Quarterly*, 38, 85–106.
- Gebhard, M., Willett, J., Pablo, J., Caicedo, J., & Piedra, A. (2011). Systemic functional linguistics, teachers' professional development, and ELLs' academic literacy practices. In

- T. Lucas (Ed.), *Teacher preparation for linguistically diverse classrooms: A resource for teacher educators* (pp. 55–73). New York, NY: Routledge.
- Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P. W. (2009). Teaching practice: A cross-professional perspective. *Teachers College Record*, 111(9), 2055–2100.
- Harper, C. A., & de Jong, E. (2004). Misconceptions about teaching English-language learners. *Journal of Adolescent & Adult Literacy*, 48(2), 152–162.
- Harper, C. A., & de Jong, E. (2009). English language teacher expertise: the elephant in the room. *Language and Education*, 23(2), 137–151.
- Jacobson, E., Remillard, J., Hoover, M., & Aaron, W. (in press). The interaction between measure design and construct development: Building validity arguments. In A. Izsák, J. T. Remillard, & J. Templin (Eds.), *Psychometric methods in mathematics education: Opportunities, challenges, and interdisciplinary collaborations* (pp. xxx–xxx). Journal for Research in Mathematics Education monograph series. Reston, VA: National Council of Teachers of Mathematics.
- Lai, Y., & Howell, H. (2014). *Tasks assessing mathematical knowledge for teaching as representations of teaching practice*. Unpublished manuscript. University of Nebraska-Lincoln & Educational Testing Service.
- Lindlof, T. R. (1995). *Qualitative communication methods*. Thousand Oaks, CA: Sage.
- Loeb, S., Soland, J., & Fox, L. (2014). Is a good teacher a good teacher for all? Comparing value-added of teachers with their English learners and non-English learners. *Educational Evaluation and Policy Analysis*, 36(4), 457–475.
- Lucas, T., & Villegas, A. M. (2011). A framework for preparing linguistically responsive teachers. In T. Lucas (Ed.), *Teacher preparation for linguistically diverse classrooms: A resource for teacher educators* (pp. 55–73). New York, NY: Routledge.
- Lucas, T., Villegas, A. M., & Freedson-Gonzalez, M. (2008). Linguistically responsive teacher education preparing classroom teachers to teach English language learners. *Journal of Teacher Education*, 59(4), 361–373.
- Martiniello, M. (2008). Language and the performance of English-language learners in math word problems. *Harvard Educational Review*, 78(2), 333–368.
- Masingila, J. O., & Doerr, H. M. (2002). Understanding pre-service teachers' emerging practices through their analyses of a multimedia case study of practice. *Journal of Mathematics Teacher Education*, 5, 235–263.
- McCarthy, M. (1991). *Discourse analysis for language teachers*. Cambridge, UK: Cambridge University Press.
- Mislevy, R. J. (1994). Evidence and inference in educational assessment. *Psychometrika*, 59, 439–483.
- Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2003). *A brief introduction to evidence-centered design* (Research Report 03-16). Princeton, NJ: Educational Testing Service.

- Moschkovich, J. N. (1999). Supporting the participation of English language learners in mathematical discussions. *For the Learning of Mathematics*, 19(1), 11–19.
- Moschkovich, J. (2002). A situated and sociocultural perspective on bilingual mathematics learners. *Mathematical Thinking and Learning*, 4(2&3), 189–212.
- National Clearinghouse for English Language Acquisition. (2011). *The growing numbers of English learner students*. Retrieved from http://www.ncela.gwu.edu/files/uploads/9/growingLEP_0809.pdf
- Oakes, J. (1990). *Multiplying inequalities: The effects of race, social class and tracking on opportunities to learn mathematics and science*. Santa Monica, CA: Rand Corporation.
- Reeves, J. R. (2009). A sociocultural perspective on ESOL teachers' linguistic knowledge for teaching. *Linguistics and Education*, 20, 109–125.
- Schleppegrell, M. J. (2001). Linguistic features of the language of schooling. *Linguistics and Education*, 12(4), 431–459.
- Schleppegrell, M. J. (2004). *The language of schooling: A functional linguistics perspective*. Mahwah, NJ: Erlbaum.
- Schleppegrell, M. J. (2013). The role of metalanguage in supporting academic language development. *Language Learning*, 63(1), 153–170.
- Secada, W. (1992). Race, ethnicity, social class, language and achievement in mathematics. In D. Grouws (Ed.), *Handbook for research on mathematics teaching and learning* (pp. 623–660). New York, NY: Macmillan.
- Secada, W. G. (1998). School mathematics for language enriched pupils. In S. H. Fraud & O. Lee (Eds.), *Creating Florida's multilingual global work force: Educational policies and practices for students learning English as a new language*. Tallahassee, FL: Florida Department of Education.
- Turkan, S., Croft, A., Bicknell, J., & Barnes, A. (2012). *Assessing quality in the teaching of content to English language learners*. (Research Report 12–10). Princeton, NJ: Educational Testing Service.
- Turkan, S., de Oliveira, L., Lee, O., & Phelps, G. (2014). Proposing the knowledge base for teaching academic content to English language learners: Disciplinary linguistic knowledge. *Teachers College Record*, 116(3), 1–30.
- U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. (2015). Retrieved from http://www.nationsreportcard.gov/reading_math_g12_2013/#/changes-by-groups

Appendix

Interview questions

1. Can you tell in your own words what information the question gives you and what it asks you to do?
2. What was your answer?
3. What information in the item did you draw on when you selected this answer choice ____ (e.g., A)?
4. What information in the item helped you to eliminate the other choices _(e.g., B, C, D)? OR
 - *You indicated that you chose __as the answer. Why do you think that's the best answer? Let's go through the other options. Why didn't you select __?*
5. Does the scenario present a challenge of teaching ELLs that you would encounter in your classroom? Which middle grade level is the scenario most appropriate?
6. Does the scenario present a learning challenge that you would encounter with your ELL(s) in your classroom?
7. On a scale of 1-5, how important do you think it is for (math or science) teachers of ELLs to identify the ELL language proficiency levels? On a scale of 1-5, how useful was it to be informed about the proficiency levels of the ELL(s) characterized in the item?
8. On a scale of 1 to 5, with 5 being the most important, how important do you think the knowledge being tested in this item is for math/science teachers who teach ELLs?
 - How important is it for teachers of ELLs to know how to answer this question?
 - How important is it for teachers of ELLs to be able to answer this question?
9. Was there anything about the question/scenario/problem that you found unclear or ambiguous?

Item 1.

Read the learning objective, teacher task, characteristics of focus ELLs, and instructional scenario, and then answer the question that follows.

Learning objective: Students will understand how to rewrite expressions using the distributive property.

Teacher task: The teacher is introducing students to the distributive property and providing examples to scaffold the students' understanding of the distributive property.

Characteristics of focus ELLs: ELLs in the class are newly arrived in the school. Their placement-test results indicate that they are weak in mathematics and have low proficiency in English.

Instructional scenario: While preparing for a lesson on the distributive property, a teacher looks up the definition given in the textbook. The definition is shown below.

Distributive property—the product of a number and a sum is equal to the sum of the individual products of the addends and the number

Rather than begin the lesson with the definition, the teacher wants to give an illustration of the distributive property.

Which of the following examples should the teacher write on the board to best scaffold the ELLs' understanding?

A. $a(b + c) = ab + ac$

B.

$$2 \times 3 + 2 \times 4$$

$$6 + 8 = 14$$

$$2 \times 7 = 14$$

$$2(3 + 4) = 2(3) + 2(4)$$

C. At a store, each book costs \$4, and each video game costs \$15. Ellen is buying a book and a video game for each of her 3 children. How much does Ellen spend altogether?

D. The ingredients in a cookie recipe include 4 cups of flour and 2 eggs. Bob is doubling the recipe. How many cups of flour and how many eggs will Bob need to make the cookies?

Item 2.

Read the learning objective, student task, characteristics of focus ELLs, and instructional scenario, and then answer the question that follows.

Learning objective: All students will understand how to write and solve equations based on word problems.

Student task: Students will write and then solve an equation based on a word problem.

Characteristic of focus ELLs: ELLs in the class have grade-level content knowledge, and their English reading proficiency has been identified as being at WIDA Reading Level 3.

Instructional scenario: During a unit on solving word problems that involve multi-step equations, a teacher gives the problem below to the students.

In 1996, the salary of the governor of New York was about \$50,000 less than triple the salary of the governor of Arkansas. The total of the two salaries was \$190,000. Find the 1996 salary of each state's governor.

Which of the following is the best strategy to support the ELLs in being able to understand and solve the problem?

- A. Provide ELLs with explanations of the difficult phrases in the problem.
- B. Provide ELLs with a visual illustration of the problem.
- C. Give ELLs a version of the problem in simplified English.
- D. Give ELLs the equation that the problem is based on and have them solve the equation.

Teachers and their Educators – Views on Contents and their Development Needs in Mathematics Teacher Education

Mika Koponen, Mervi A. Asikainen, Antti Viholainen and Pekka E. Hirvonen
University of Eastern Finland

Abstract: Finland has scored well in international assessments (e.g. PISA, TIMSS), and the pressure to attain excellent scores has activated a drive toward even more effective mathematics teacher education. This article presents the results of a qualitative assessment of the mathematics teacher education provided by the University of Eastern Finland. In this study, the views held by practicing teachers (N=101) and teacher educators (N=19) are compared so that the outstanding development needs of mathematics teacher education in terms of their contents can be revealed. The data was gathered via an electronic survey and was mainly analyzed using data-driven methods. In addition, framework provided by Mathematical Knowledge for Teaching (MKT) was used to categorize the respondents' views regarding the contents of mathematics teacher education and to develop general guidelines for the reform of mathematics teacher education. The results indicate that mathematics teacher education should include pure mathematical content (Common Content Knowledge, CCK) and mathematical content that will have been designed only for future teachers (Specialized Content Knowledge, SCK). Teacher educators and practicing teachers both held the view that the relevance of CCK studies depend on the connections between university and school mathematics. Pedagogical studies should also be reformed because practicing teachers have realized that effective teaching (Knowledge of Content and Teaching, KCT) requires knowledge about learning mathematics (Knowledge of Content and Students, KCS) that is not offered in the current educational system on a sufficiently broad basis. In this study, suggestions for developing mathematics teacher education were mostly connected to four domains of MKT: (CCK, SCK, KCT and KCS). Interestingly those domains are the same domains which has been empirically tested and better conceptualized.

Keywords: Mathematical Knowledge for Teaching, MKT, mathematics teacher education, evaluating teacher education, contents of mathematics teacher education.

Introduction

Finland has scored well in international assessments (e.g., PISA, TIMSS), and the Finnish school system has been rated as being of top quality. Finnish teacher education has also been evaluated as high in quality from an international perspective (Kivirauma & Ruoho, 2007; Tryggvason, 2009). An important reason for this success is that Finnish teachers are educated both systematically and extensively, and every qualified teacher must have a Master's degree (Tryggvason, 2009). It is claimed that Finnish teacher education is the result of a long-term, research-based development (Tryggvason, 2009). However, the voices of practicing mathematics teachers and teacher educators have not received attention enough in the research field. Are these two groups satisfied with the current contents of mathematics teacher education and what kind of needs of development they see at the moment?

In the present study we focus on practicing mathematics teachers' and teacher educators' views on mathematics teacher education. The practicing teachers participating in this study graduated in the period of 2002–2012 and they nowadays teach at school level, which enables them to evaluate the contents of teacher education from a perspective of the teacher's profession. In addition, when the survey was implemented the teacher educators were actively working as teacher educators. We were interested in discovering how these two subject groups saw the present contents

of the Mathematics Teacher Education Program (MTEP) at the University of Eastern Finland and also in how they would develop the teacher education program. We sought answers to the following research questions:

1. How do teacher educators and practicing mathematics teachers regard the course contents of mathematics teacher education?
2. What kind of recommendations would teacher educators and practicing mathematics teachers make for improving mathematics teacher education program?

The views held by practicing teachers and teacher educators play an import role in developing teacher education. There may be a possibility that the contents are not regarded as being as useful as teacher educators assume. It is also possible that practicing teachers and teacher educators hold conflicting views about the contents. Hence, the views of both groups are important in order to be able to form a coherent picture of the current status of teacher education and to construct an extensive basis for the development work.

Our methodical aim has been to test a theoretical framework called *Mathematical Knowledge for Teaching (MKT)* (Ball, Thames & Phelps, 2008) through the process of categorizing practicing teachers' and teacher educators' views. This framework appeared to be promising for categorizing these views, since it has previously worked relatively well in classifying teacher knowledge (see Markworth, Goodwin, & Glisson, 2009; Fauskanger, Jakobsen, Mosvold, & Bjuland, 2012).

Conceptualizing the Teaching of Mathematics

Mathematical knowledge for teaching

There was an increasing interest in the 1980s in teacher qualifications and methods of effective teaching that would influence student learning. Lee Shulman proposed that a teacher also needs to possess other types of knowledge than pure subject matter knowledge in order to teach so that students would understand. In 1986 Lee Shulman introduced a new term, *pedagogical content knowledge (PCK)*. According to Shulman (1986), teachers must have an integrated knowledge of subject and pedagogy, some kind of amalgam knowledge. Initially, Shulman considered PCK to be a topic-specific subcategory of content knowledge, which included two further subcategories: knowledge of representations and knowledge of learning difficulties and strategies for overcoming them. Shulman's later model consisted of seven categories, of which PCK was one, with no subcategories (Shulman, 1987). By proposing PCK as one out of seven categories of conceptualization, Shulman neglected the potential for integration among these categories and the hierarchies that might exist between them, and left the task of further development of the concept to other researchers (Hashweh, 2005).

Shulman's conceptualization has been criticized for its restricted and ambiguous definitions of categories (Ball et al., 2008, Hashweh, 2005). Ball et al. (2008) claim that the terms PCK and *content knowledge* are frequently confused with common pedagogical skills. Meredith (1995) argues that PCK as defined by Shulman simply implies one type of pedagogy rooted in particular representations of prior knowledge. Meredith suggests that learners have a built-in competence for constructing their own understanding of subject matter, but Shulman's PCK seems not to encompass alternative views of teaching. Meredith argues that Shulman's definition of PCK leads to teaching methods where the teacher will explain and illustrate procedures while learners practice the procedures by using examples. Thus, the teacher's role can be seen as transmitting mathematical knowledge and helping learners to acquire understanding.

Shulman's conceptualization has also been claimed to ignore the interaction between the different categories, assuming that knowledge is static rather than possessing a dynamic nature (Hashweh, 2005; Fennema & Franke, 1992). Fennema and Franke (1992) argue that teacher

knowledge frequently changes in light of classroom interaction experiences, and hence teachers' beliefs should form an important part of the conceptualization. According to Fennema and Franke, teacher knowledge can be divided into four parts: *knowledge of content*, *knowledge of pedagogy*, *knowledge of students' cognitions*, and *teachers' beliefs*. At the center of this model is *context specific knowledge*, which can be seen as dynamic knowledge, since it occurs in the context of the classroom. In this model, PCK consists of teachers' knowledge of teaching procedures, such as effective strategies for planning, classroom routines, behavior management techniques, classroom organization procedures, and motivational techniques. Fennema and Franke (1992) see teacher knowledge as interactive and dynamic in nature and they suggest that no single domain of teacher knowledge plays a particular role in the effective teaching of mathematics.

Rowland, Turner, Thwaites & Huckstep (2009) developed *The Knowledge Quartet* conceptualization, which was based on Shulman's conceptualization (1986) with respect to Fenneman and Franke conceptualization (1992). The Knowledge Quartet was generated by categorizing elementary teachers' classroom actions. The main aim of the research work was to investigate the relation between the teacher's subject matter and PCK knowledge. Detailed analysis of the elementary mathematics lessons taught by pre-service teachers resulted in the identification of teacher knowledge framework. Rowland et al. (2009) suggest that the framework can be used to classify teachers' actions in the context of a classroom.

One of the most promising recent efforts in discovering the kind of knowledge and skills that are needed for high-quality mathematics teaching has been the theoretical framework known as *Mathematical knowledge for teaching (MKT)*, as posited by Ball and her associates¹. In this model, subject matter knowledge is categorized into three domains: *common content knowledge (CCK)*, *horizon content knowledge (HCK)*, and *specialized content knowledge (SCK)* (see Figure 1). In addition, PCK consists of three parts: *knowledge of content and student (KCS)*, *knowledge of content and teaching (KCT)*, and *knowledge of content and curriculum (KCC)*. The domains CCK, HCK, and SCK are subject matter knowledge that requires no knowledge concerning either the students or pedagogy. In addition, the domains of KCS, KCT, and KCC are the kind of knowledge that requires an integrated knowledge made up of subject matter knowledge and pedagogical knowledge (Sleep, 2009), as in Shulman's (1986) conceptualization. According to Sleep (2009), four of the domains (CCK, SCK, KCS and KCT) have been empirically tested and better conceptualized, while two of the domains (HCK and KCC) are still in the earlier stages of conceptualization.

¹The University of Michigan projects *Mathematics Teaching and Learning to Teach project (MTLT)* and *Learning Mathematics for Teaching project (LMT)* produced plenty of details to form MKT, e.g., Hill & Ball, 2004; Hill, Schilling & Ball, 2004; Hill, Rowan & Ball, 2005; Hill & Lubienski, 2007; Hill, Ball, Sleep et al., 2007; Hill, 2007; Schilling, 2007; Schilling, Blunk & Hill, 2007; Schilling & Hill, 2007; Hill, Ball, Blunk, et al., 2007; Hill, Dean & Goffney, 2007; Hill, Ball & Schilling, 2008; Delaney, Ball, Hill et al., 2008; Stylianides & Ball, 2008; Hill, Blunk, Charalambous et al., 2008; Ball, Thames & Phelps, 2008; Ball & Forzani, 2009; Thames & Ball, 2010.

Figure 1. Domains of Mathematical Knowledge for Teaching (MKT) by Ball et al. (2008)

CCK consists of mathematical knowledge and skills used in any settings, including in settings other than teaching, and it includes calculating, solving problems, and other common mathematical knowledge that is not unique to teaching (Ball et al., 2008). SCK is mathematical knowledge and skills that are peculiar to teaching, and is typically not intended for other settings than teaching (Ball et al., 2008). In other words, SCK consists of the mathematical knowledge and skills that a mathematician does not need, while at the same time they are needed by a teacher in order to practice effective teaching. HCK consists of mathematical knowledge of the mathematical structures and also awareness of how mathematical topics are related to each other in a curriculum (Ball et al., 2008). This means that a teacher needs to know how topics are related to each other at different school levels and how mathematics is actually constructed.

KCS consists of amalgam knowledge of students, learning, and mathematics (Ball et al., 2008). A teacher must be able to anticipate students' difficulties, hear and respond to students' thinking, and choose suitable examples and presentations while teaching. A teacher's action in planning and teaching requires awareness of students' conceptions and misconceptions of different mathematical topics. KCT is also amalgam knowledge of teaching and mathematics (Ball et al., 2008). Teachers need KCT knowledge in choosing proper activities, exercises and representations for different topics. Teachers need KCT knowledge for both planning and teaching. One important part of this knowledge for teachers is to recognize situations where teachers should diverge from their original planning, for example, if a student makes a mathematical discovery.

KCC represents amalgam knowledge of mathematics and curriculum. According to Sleep (2009), a teacher needs to know the contents of the curriculum, but Ball et al. (2008) offer only a restricted definition of KCC and hence the kind of knowledge and skills that KCC includes remains unclear. Our preliminary analysis of the data in the present study showed that if MKT is used to organize practicing teachers' and educators' views, the KCC domain has to be modified. The practicing teachers and teacher educators mentioned skills and knowledge related to teaching equipment. Hence, our conceptualization states that KCC also includes knowledge and skills related to teaching materials (including textbooks, other materials, etc.), teaching instruments (blackboard, overhead projector, etc.), and technology (computer, smart board, calculators, software, etc.).

The Evolution of MKT

The development of MKT started with the study of classroom actions with a view to identifying the knowledge needed for teaching mathematics (Ball & Bass, 2003). This work continued with the formation of hypothetical characterizations of MKT (e.g. Ball, Hill & Bass, 2005; Ball et al., 2008). Thereafter, Hill, Schilling, and Ball (2004) developed specific measurements of MKT that could be used to test this hypothetical characterization. In the case of validating measurements, the Michigan group tested measurements against practice (Hill, Blunk, Charalambos, et al., 2008) and also against students' achievements (Hill, Rowan & Ball, 2005). Thereafter, MKT has been used to develop the contents of teacher education in ways that should help teachers to acquire the knowledge required for teaching mathematics (Ball, Sleep, Boerst & Bass, 2009).

Markworth, Goodwin and Glisson (2009) have used MKT to evaluate what student teachers have learned during a teaching practicum course. They coded interview responses and conversational topics on the basis of the domains of MKT. By using MKT in their analysis, Markworth, Goodwin and Glisson (2009) were able to capture more detailed information about the subject matter knowledge and pedagogical content knowledge that student teachers had gained during the teaching practicum course.

In the course of this study, practicing teachers and teacher educators suggested various recommendations for improving the mathematics teacher education program. To identify these suggestions systematically, we used MKT in a similar way to that of Markworth, Goodwin and Glisson (2009). This meant that suggested recommendations for improving mathematics teacher education could be classified in terms of six domains of MKT.

Method

Context

This study was implemented at the University of Eastern Finland, which offers two programs for students of mathematics: one for mathematicians and another for teachers. The programs are almost identical in their respective amounts of mathematics courses, but they differ in minor subjects. In the present study, we concentrate on the program for teachers, Mathematics Teacher Education Program, MTEP.

MTEP includes a Bachelor's degree (180 cp²) and a Master's degree (120 cp). Both degrees are required for a student to qualify as a mathematics teacher in Finland. MTEP includes mathematical studies (130 cp), pedagogical studies (60 cp), and studies in one or two minor subjects (60 cp each). Most mathematical studies are traditional mathematics courses, which are compulsory for both future teachers and mathematicians (e.g. *calculus, analysis, algebra, differential equations*, etc.).

The pedagogical studies include theoretical studies focusing on teaching and learning (30 cp), the didactics of mathematics (10 cp), and teaching practice (20 cp). Teaching and learning courses are intended to all subject teachers and courses are concerning teaching and learning in general. However, the following courses which are intended only to forthcoming subject teachers of mathematics enables taking into account the special aspects of mathematics. Teaching practice is undertaken at the university teacher training school. Student teachers plan their own teaching sequences or lessons under the guidance of a subject teacher. Student teachers' lessons are evaluated and feedback is also provided. The amount of student teaching is approximately 50

² One *credit point* (cp) is the equivalent of 25 hours of study. The recommendation is to complete 60 cp of studies per year.

lessons. The training school teachers' task is to guide student teachers in addition to performing their own ordinary teaching work.

Student teachers can choose to study any school subject as a minor subject, but the most typical choices are physics or chemistry, or both. In its entirety, MTEP provides students with the competence to teach mathematics and minor subjects at lower or upper secondary schools and vocational schools.

Sample

The data was collected in the course of two separate electronic surveys conducted in 2012–2013. The first survey was aimed at mathematics teachers who had graduated from the UEF during the period 2002–2012. Our sample (N=101) includes 54% of all teachers who have graduated from UEF during period 2002–2012. In the sample, the majors taken by our respondents were 72% (73) mathematics, 20% (20) physics and 8% (8) chemistry, which makes the sample similar to the distribution of graduated teachers according to their major subject. All of the respondents, with the exception of one, had had previous experience of teaching mathematics at school or they were working as teachers when the survey was implemented.

The second survey targeted the teacher educators in mathematics at the UEF, who taught either mathematical studies or pedagogical studies or were guiding teaching practice. Our sample (N=19) includes 79% of all of the teacher educators in mathematics at the UEF. In the sample, 74% (14) of the teacher educators taught mathematics and 26% (5) worked in pedagogical studies or teaching practice. To conceal the respondents' identities, the teacher educators in the fields of both pedagogical studies and teaching practice were placed in the same category.

Instrument and data analysis

The study was implemented with the aid of a survey that included statements about the knowledge and skills learned in MTEP and open questions about the present state and the future of MTEP.

The survey conducted with practicing teachers included three open questions about MTEP.

1. Evaluate the contents of mathematical studies in MTEP, especially with regard to the work of a mathematics teacher.
2. Evaluate the contents of the pedagogical studies and teaching practice in MTEP, especially with regard to the work of a mathematics teacher.
3. Suggestions regarding the development of mathematics teacher education would also be appreciated.

The survey involving the UEF teacher educators included two open questions.

4. Evaluate the contents of the studies you teach, especially with regard to the work of a mathematics teacher.
5. Please make a suggestion regarding the development of mathematics teacher education would also be appreciated.

The data was analyzed using qualitative content analysis (Tesch, 1990; Hickey & Kipping, 1996; Mayring, 2000; Hsieh & Shannon, 2005). Hsieh and Shannon (2005) have identified three different approaches to qualitative content analysis that can be used to interpret meaning from the content of text data: conventional, directed or summative (Hsieh & Shannon, 2005).

Our analysis started with reading the data several times to achieve immersion and obtain a sense of the whole (Tesch, 1990). Then, practicing teachers' and teacher educators' perceptions about the contents of MTEP (Questions 1, 2 and 4) were analyzed with *Conventional Content Analysis* (Hsieh & Shannon, 2005). In the conventional content analysis, coding categories are

derived directly from the text data. In our data, respondents' personal experience or more like attitudes towards contents emerged clearly from data. Each respondent was placed in one of these categories (Figure 2).

A majority of the practicing teachers' mentioned only issues that should be developed in the contents of MTEP (Questions 1 and 2), and therefore their responses were placed in the category of *In need of development*. The contents of this category were analyzed with directed content analysis, which is a more structured process than the conventional approach (Hickey & Kipping, 1996; Hsieh & Shannon, 2005). Direct content analysis starts with a theory, which is used for coding text data (Hsieh & Shannon, 2005). Generally, a goal of the directed approach is to validate or conceptually extend a theoretical framework or theory (Hsieh & Shannon, 2005). *Mathematical Knowledge for Teaching (MKT)* framework was a starting point for designing the survey, and each statement was designed to be interconnected to the domain of MKT. In the planning, we noticed a possibility for using MKT for directed content analysis. A pre-analysis of the data indicated that all the issues in question 1 and many of the issues in question 2 can be categorized with MKT. Issues beyond MKT were categorized with the conventional content analysis in case of question 2.

Both the surveys also included blank spaces for other suggestions for the development of the teacher education program (Questions 3 and 5). Many of the respondents did, however, mention the same issues which they already mentioned in the previous question related to the contents. Therefore we used directed content analysis similarly as in the categorization of the suggestions related to the six domains of MKT. Suggestions beyond MKT were categorized with the conventional content analysis. Previous questions in the survey covered the majority of respondents' ideas, and so there were only a few new ideas among these suggestions.

Figure 2. Text data analysis was performed with conventional and direct content analysis (Hsieh & Shannon, 2005).

Results

The results of the study are presented in two parts. First, we discuss how teacher educators and practicing mathematics teachers view the contents of mathematics teacher education. Second, our discussion focuses on teacher educators' and practicing mathematics teachers' ideas for developing teacher education. Suggestions for developing mathematics teacher education will be represented in tables where the categories have been provided mainly by MKT.

Views on the contents of mathematics teacher education

Practicing teachers' views concerning the contents of mathematics studies. The categorization of practicing mathematics teachers' views concerning the contents of mathematical showed that one fifth of the respondents (21%) viewed the contents neutrally. Half of them considered the number of mathematics courses appropriate for teachers, while the other half gave no reasons for their responses. A small minority of the respondents (7%) did not consider the contents of the mathematics courses useful for teachers. In most cases, the reason for this was that the courses were considered to provide too complex a discussion of mathematics in comparison with the mathematics needed in a teacher's work. No fully positive views appeared in the categorized responses.

A majority of the practicing teachers (59%) provided only suggestions related to developing the present contents of mathematics courses. These suggestions were analyzed again by using MKT. Most of these suggestions (79%) were related to improving student teachers' subject matter knowledge, while one fifth of them were concerned with developing student teachers' pedagogical knowledge and skills (see Table 1).

Table 1. *Categorization of practicing teachers' (N=60) suggestions for developing the content of mathematical studies. Each respondent was permitted to mention more than one issue.*

Category	Domain of MKT	f
Subject matter knowledge and skills	<i>Common content knowledge (CCK)</i>	
	• Present course contents are not linked with school mathematics	12
	• Present course contents are not the same as in schools	11
	• More geometry	4
	• More financial and statistical mathematics	2
	• Wider knowledge of mathematical concepts	2
	<i>Specialized content knowledge (SCK)</i>	
	• Present mathematical studies should be separate for student teachers and mathematicians	20
	• More school mathematics needed	12
	<i>Horizon content knowledge (HCK)</i>	
Pedagogical content knowledge and skills	• Present course contents are not linked with each other	4
	• More skills concerned with teaching students at different levels	1
	<i>Knowledge of content and students (KCS)</i>	
	• More studies concerning learning difficulties in mathematics	4
	• More skills concerned with teaching students at different levels	3
	<i>Knowledge of content and teaching (KCT)</i>	
	• More courses about didactic mathematics	6
	• More courses about how to differentiate teaching	2
	• More skills to motivate students in mathematics	1
	• More studies about teaching problem solving	1
	<i>Knowledge of content and curriculum (KCC)</i>	
	• More courses about using technology in teaching mathematics	1

Common content knowledge (CCK). The practicing teachers mentioned that the present contents of mathematical studies are not the same as the contents of school mathematics. They were disappointed that they had studied so much mathematics that they had never used in their school teaching. The practicing teachers also mentioned that the present contents did not link properly with school mathematics. Some teachers claimed that presentations at university were either symbolic and theoretical or too complex in comparison with school mathematics, and hence it was hard to see how the course contents were linked with school mathematics. The practicing teachers mentioned that they lacked the competence to teach geometry and financial or statistical mathematics, and so they suggested that MTEP should include more courses in those domains. It is evident that practicing teachers need mathematical content knowledge (Ball et al., 2008), but the opinions of the practicing teachers indicated that to be useful for future teachers, the contents should be linked to school mathematics.

Specialized content knowledge (SCK). The practicing teachers suggested that mathematical studies should be arranged differently for future teachers and for mathematicians. They argued that the current integrated mathematical courses do not support future teachers properly. Some practicing teachers said that at university the focus of mathematics was proving and presenting results, whereas school mathematics consisted of rather more than that. Almost all the respondents suggested that the mathematical representations of the course contents should be modified with respect to teachers' actual work. According to them, in the current situation MTEP includes too much pure mathematics and not enough school mathematics. Most of them recalled that in MTEP there was a course called *School mathematics*, which they found important and useful. The contents of the course were the same as in actual school mathematics, and the implementation of the course resembled the mathematics teaching conducted in schools. In consequence, they argued that they learned the contents of such courses well and that they had been able to use the course contents in their teaching work. All of them argued that there should be more courses of this kind in MTEP. All

of these practicing teachers' views are linked to the definition of SCK (Ball et al., 2008): practicing teachers need mathematical knowledge that is particular to the needs of teachers.

Horizontal content knowledge (HCK). The practicing teachers argued that the contents of university mathematics courses were not interconnected or that the links could not be detected during the courses. In their view, courses that were in fact extensions of each other (e.g., calculus 1, calculus 2) were separate courses; alternatively, they were unable to detect the ways in which new mathematical concepts could be constructed on the basis of previously learned concepts. In the view of the respondents, the mathematical knowledge base ought to resemble a network, while, for them, the contents of MTEP did not support the construction of that kind of concept. Mathematical knowledge lacking a proper understanding of the structure of mathematics can be identified as the major challenge in the domain of HCK (see Ball & Bass, 2009).

Knowledge of content and students (KCS). The practicing teachers claimed that pedagogical issues can also be discussed during a mathematics course. They mentioned that they did not develop any clear idea of how students were actually learning mathematics during the mathematics courses. Some practicing teachers mentioned that they had too little competence in the issues concerned with mathematical learning difficulties. Some teachers also argued that they need more skills related to teaching mathematics to both weak and talented students at the same time.

Knowledge of content and teaching (KCT). The practicing teachers argued that issues concerned with teaching mathematics can also be handled in mathematical studies. They mentioned that didactic mathematics and studies about how to differentiate teaching should be included in mathematical studies.

Knowledge of content and curriculum (KCC). One practicing teacher argued that future mathematics teachers needed more enhanced skills concerned with the use of technology in teaching mathematics since teachers were increasingly using technology in schools.

Practicing teachers' views about the contents of pedagogical studies and teaching practice. The categorization of practicing mathematics teachers' views about the contents of pedagogical studies and teaching practice demonstrated that the respondents' views were diverse. Small minorities of the respondents viewed these studies positively (2%) as useful for teachers; or neutrally (9%), often without providing reasons; or negatively (7%), considering the courses useless for teachers; but most of them (67%) consider that there was a need for development in these studies.³ The categorization of their suggestions is presented in Table 2. These suggestions mainly concerned pedagogical content knowledge and skills (51%), and development of the structure of mathematics teacher education (40%).

³ 12% of respondents did not answer this question, and the responses of 3% were irrelevant.

Table 2. *Categorization of practicing teachers' (N=68) views on how to develop the content of pedagogical studies and teaching practice. Each respondent was permitted to mention more than one issue.*

Category	Domain of MKT	f
Pedagogical content knowledge and skills	<i>Knowledge of content and students (KCS)</i>	
	• More studies of learning difficulties in mathematics	13
	• More skills concerned with how to handle students at different levels	10
	• More skills concerned with evaluating students' knowledge and skills	6
	• More studies concerned with other learning theories	2
	<i>Knowledge of content and teaching (KCT)</i>	
	• More studies of teaching mathematics; didactic mathematics	10
	• More training in the planning and teaching of complete courses	6
	• More studies of how to differentiate teaching	5
	• More skills concerned with motivating students of mathematics	4
	• More courses about functional teaching methods, teaching problem-solving, or visualizing mathematics	3
	• More studies of how to link learning theories to practice	1
	<i>Knowledge of content and curriculum (KCC)</i>	
	• More skills and knowledge to produce teaching materials of their own	4
	• More studies of using technology in teaching mathematics	3
Structure of mathematics teacher education	<i>Amount of studies</i>	
	• More teaching practice	21
	• More pedagogical studies	5
	• Compulsory update education after some years of teaching	1
	• More studies of how to teach minor subjects	1
	<i>Quality of studies</i>	
	• Linking theory to practice	25
	• Educators of practice teachers should give more advice about didactic issues	1
	<i>Developing curriculum of MTEP</i>	
	• Contents of pedagogical studies and practice should be better integrated	1
General issues	<i>The other knowledge and skills</i>	
	• More studies of teachers' extramural duties	10
	<i>Common issues</i>	
	• Departments' cooperation should be improved	1

Knowledge of content and students (KCS). The practicing teachers suggested that the pedagogical studies and teaching practice should include more courses about the learning difficulties encountered in mathematics. Some teachers said that they were struggling with students who probably had learning difficulties and hence they needed more skills in order to be able to recognize and handle such students. Some of the practicing teachers also mentioned that the skills concerned with teaching students at different levels would be useful for them because student groups were often very heterogeneous. The practicing teachers also mentioned that they needed more knowledge and skills for evaluating student learning and more knowledge concerned with various learning theories, since students learn in different ways.

Knowledge of content and teaching (KCT). The practicing teachers demanded more skills for teaching mathematics. Some of them mentioned that the courses in didactic mathematics were useful for them, but that they needed more knowledge of this kind. The practicing teachers argued that their studies did not include enough courses on planning and teaching complete courses. They stressed that planning was the first thing that new teachers needed to undertake after graduation. The practicing teachers also mentioned that more skills for differentiating teaching and increasing student motivation would be of assistance.

Knowledge of content and curriculum (KCC). The practicing teachers said that textbooks or other printed material did not always fit their ideas about teaching, and so they would need more

skills to design their own teaching materials. The practicing teachers also mentioned that they needed more knowledge and skills concerned with using technology in teaching mathematics in a pedagogically reasonable way.

The number of courses. The practicing teachers said that both the teaching practice and the pedagogical studies were very useful and suggested that their number should be increased in MTEP. They felt that the teaching practice was a good place for trying out new teaching methods or for trying to transform pedagogical knowledge into practice. They also told about the use of useful and functional teaching methods learnt during their teaching practice in their actual work. Some of them mentioned encountering similar situations in the classrooms to those that had been discussed in the pedagogical studies, which had helped them to better understand the relevance of the pedagogical studies.

The quality of courses. The practicing teachers argued that the courses in the pedagogical studies and even courses about teaching and learning were too theoretical, which made linking theory with practice difficult. They described a feeling of learning a lot during these studies, but without having the necessary skills to apply this knowledge in classroom situations. Some teachers even felt that the pedagogical studies were useless because they had too few links with real-life teaching situations.

Other knowledge and skills. Some practicing teachers said that they were surprised by the duties that teachers had outside the classroom. They suggested that these issues should be discussed in mathematics teacher education.

Teacher educators' views on the contents of their own courses. One fourth of the teacher educators (26%) viewed their own courses positively and considered that the courses were useful for future teachers (see Table 3). Many of them (42%) viewed their own courses neutrally and regarded the courses as having been only partly useful. Some teacher educators (16%) viewed their own courses negatively and indicated problems in the contents of courses that made them not very useful for teachers.

Table 3. *Teacher educators' (N=19) views about their courses and their suitability for future mathematics teachers. ME = teacher educator in mathematical studies, PTE = teacher educator in pedagogical studies and teaching practice.*

Class and justification	ME (N=14)	PTE (N=5)
<i>Positive 26% (5)</i>		
• Contents increase pure mathematical knowledge and the teaching methods used teach how to teach mathematics	2	0
• Contents are the same as in school mathematics	1	0
• No justification	0	2
<i>Neutral 42% (8)</i>		
• Contents are not the same as in school mathematics, but studies develop mathematical thinking	2	1
• Only some parts of contents link with school mathematics	2	0
• Some contents go beyond school mathematics or general knowledge for teachers	2	0
• Courses are non-compulsory for student teachers and therefore their contents are not useful for teachers	1	0
<i>Negative 16% (3)</i>		
• Some courses are simply all-around education for teachers and in some courses there is not enough time to teach important issues	1	0
• Students' knowledge is poor at the beginning of courses and therefore they cannot learn the contents	0	1
• Teachers learn contents but they have insufficient skills for using this knowledge in school teaching	1	0
<i>Empty 16%(3)</i>	2	1

Positive (5). There were three mathematics educators (MEs) and two pedagogical studies and teaching practice educators (PTEs) who considered that the contents of their own course were useful for teachers. The PTEs did not justify their views, but the MEs argued that the contents of their courses increased student teachers' mathematical knowledge and the teaching methods modeled the way to teach mathematics. The MEs underlined the significance of presenting things; they argued that it was important for student learning that the educator demonstrated how things worked. One ME argued that the contents of his courses were the same as in school mathematics and hence the contents were useful for future teachers.

Neutral (8). Seven MEs and one PTE considered the contents of their courses only partly useful for future teachers. One PT and two MEs claimed that, despite the contents not being the same as for school mathematics, the courses nevertheless developed students' mathematical thinking, which was also important for future teachers. Two MEs claimed that only some parts of the contents were linked with school mathematics and therefore these parts were useful for future teachers. Another ME thought that the contents are "good to know", but were unnecessary for teachers. Two MEs argued that the contents of their courses offered teachers only general knowledge since the contents were not specialized for use by teachers or the contents went beyond school mathematics. Both of them justified their views with the argument that future teachers needed a wide knowledge base in mathematics.

Negative (3). Two MEs and one PTE argued that the contents of their courses did not fully support future mathematics teachers. One ME claimed that in some courses s/he had too little time to teach issues that were important for teachers, while in other courses the contents were simply general knowledge for teachers. Another mathematics ME claimed that teachers usually learned the contents of his/her courses, but the course nevertheless did not provide them with the competence to apply this knowledge in their own teaching. One PE claimed that students had acquired insufficient earlier knowledge to learn the contents of his/her courses.

Teacher educators' and practicing teachers' suggestions for developing mathematics teacher education

The second research question was concerned with how practicing mathematics teachers and mathematics teacher educators would develop mathematics teacher education.

The practicing mathematics teachers made numerous suggestions for developing mathematics teacher education. The categorization of the suggestions in Table 4 shows that it would be valuable to develop teacher education both at the general level and also in terms of supporting future mathematics teachers' subject matter knowledge and pedagogical knowledge and skills. More than half of the suggestions (60%) concerned the contents of teacher education that could be categorized with MKT. One third of the suggestions (28%) focused on the quality of the teaching or the quantity of the studies that were categorized as ideas for developing teacher education program. A minority of the suggestions (12%) concerned a number of general issues related to teacher education.

Suggestions for improving the contents of teacher education mostly concerned pedagogical knowledge and skills. Practicing teachers suggested that they would add courses about learning difficulties in mathematics, the evaluation of students' mathematical know-how, and how to teach students with different levels of mathematical knowledge and skills. The practicing teachers also hoped that differentiating mathematics teaching would be discussed during teacher education, since classroom situations required that kind of competence from a teacher. They also suggested that the learning theories courses should be modified so that they would become easily applicable to one's own teaching. The practicing teachers would also add future teachers' knowledge and skills related to using technology in teaching mathematics because technology was assuming a more important role both in the classrooms and in society. Almost all of the suggestions concerning subject matter knowledge dealt with separate mathematics studies programs for future mathematicians and teachers. One common argument was that future teachers needed a different kind of mathematical knowledge from that used by mathematicians.

The ideas that were presented regarding development of the teacher education program concerned both the quality of teaching and the quantity of studies. The practicing teachers thought that the quality of teacher education could be increased by improving students' learning. This could be achieved by modifying present teaching methods as well applying new interactive teaching methods that would include discussions. The practicing teachers also argued that the studies should be modified to be less theoretical because they felt that the present studies were too theoretical, causing the students problems in understanding the course contents to any depth. The practicing teachers also thought that the teaching practice supported their teacher growth. However, they considered that the length of the teaching practice could be increased.

Table 4. *Categorization of practicing mathematics teachers' (N=101) suggestions for developing mathematics teacher education. Each respondent was permitted to mention more than one issue.*

Category	Suggestions for improving education	f
Pedagogical knowledge and skills	<i>Knowledge of contents and students (KCS)</i>	
	• Courses about learning difficulties in mathematics	7
	• How to test students' knowledge and skills	3
	• How to teach students at different stages of learning	2
	<i>Knowledge of contents and teaching (KCT)</i>	
	• How to differentiate teaching	5
	• How to bridge the gap between learning theories and practice	4
	• How to produce and use one's own teaching materials	2
	• Functional learning methods	1
	• Learner-centered teaching methods	1
	• Special education in mathematics	1
	• How to plan and teach complete courses	1
	<i>Knowledge of contents and curriculum (KCC)</i>	
Subject matter knowledge and skills	• How to use technology in teaching mathematics	6
	• Curricular knowledge	1
	• Teaching methods based on technology	1
	<i>Specialized content knowledge (SCK)</i>	
	• Separate mathematics courses for teachers and mathematicians	16
	• Problem-solving	1
	• Mathematics in different professions	1
Ideas for developing the teacher education program	<i>Horizon content knowledge (HCK)</i>	
	• Mathematical concepts at different school levels	1
	• The structure of mathematics	1
	<i>Quality of teaching</i>	
	• New teaching methods (e.g., more discussion)	10
	• Less theory – more practice – linking theory and practice	6
	• Integrating lectures and doing exercises in mathematical courses	1
	<i>Quantity of studies</i>	
	• More teaching practice	7
	• More pedagogical courses	1
	• More teaching practice and less pedagogical studies	1
General issues	<i>Improving cooperation</i>	
	• Cooperation between different departments	4
	• Cooperation between university and schools	1
	• Cooperation between students and educators; paying attention to students' suggestions regarding development	1
	<i>Special suggestions</i>	
	• Teachers should be specialized in teaching at different school levels	1
	• Subject teachers' major should be in education	1
	• Compulsory updating of education after some years of work experience	1
	<i>Beyond MKT knowledge</i>	
	• More knowledge about teachers' duties out of class	2
	<i>Uncategorized responses</i>	
	• Irrelevant	4
	• Blank	27

The teacher educators saw less reason for development than did the practicing teachers. Most of the teacher educators suggested developing mathematics teacher education by improving student teachers' subject matter studies (see Table 5).

Table 5. *Categorization of teacher educators' (N=19) suggestions for developing mathematics teacher education.*

Category	Suggestion for improving education	f
Subject matter knowledge and skills	<i>Specialized content knowledge (SCK)</i>	
	• Contents of mathematics courses should be revised to be useful for future teachers	4
	• Mathematics courses should be separately designed for teachers and mathematicians	3
	<i>Common content knowledge (CCK)</i>	
	• More courses in mathematics	2
Developing the teacher education program	<i>Horizon content knowledge (HCK)</i>	
	• New course on the structures of mathematics	1
	<i>Updating structure of studies</i>	
	• Combining mathematics and pedagogics courses as an integrated unit	1
	• Re-scheduling courses in mathematics, pedagogics, and teaching practice	1
	<i>Uncategorized responses</i>	
	• Irrelevant	2
	• Blank	5

Most of the suggestions concerned modifying future teachers' mathematical studies. Some educators suggested that the contents of present courses should be revised from the viewpoint of teacher's work and current school curricula. Many of the respondents would develop courses to increase future teachers' specialized content knowledge (SCK). Some educators also suggested that mathematics studies should be separately designed for future teachers and mathematicians, an idea that was also put forward by the practicing teachers. However, a few teacher educators argued that pure mathematics was the basis of good teaching and therefore the quantity of pure mathematics studies should be increased. One educator suggested that there was a need for developing a mathematics course whose rationale would be to link together the various domains of mathematics. We categorized this as an example of improving students' horizon content knowledge.

Two teacher educators considered that the structure of the mathematics teacher education should be updated. Another educator suggested that mathematical and pedagogical studies should not be organized separately, since their separation prevented the possibility of linking theory and practice. Another teacher educator argued that studies should be better scheduled to help student teachers to acquire an integrated knowledge of subject matter and pedagogy. It should also be noted that a third of the respondents provided no suggestions for developing mathematics teacher education, and hence it remains unknown whether these respondents were satisfied with the current teacher education or not.

Discussion

This study has investigated teacher educators' and practicing mathematics teachers' views of the contents and the development needs of mathematics teacher education as provided by the University of Eastern Finland. Practicing teachers and teacher educators made various recommendations for improving mathematics teacher education program. We consider that we have been able to identify systematically and in a detailed way the kind of subject matter knowledge and pedagogical content knowledge that these recommendations concern by classifying them in terms of the domains of MKT. Challenges concerning the content of mathematics teacher education seem to become more explicit when subject matter knowledge and pedagogical content knowledge are divided into more detailed components. Markworth, Goodwin and Glisson (2009) found similar benefits when they used MKT to evaluate a single course in mathematics teacher education. The combined results show that a majority of the recommendations concerning the issues that will need to be examined in mathematics teacher education are closely related to four domains of Mathematical Knowledge for Teaching (CCK, SCK, KCS, and KCT). Interestingly, these four

domains have been more empirically tested and better conceptualized than the other two domains (Sleep, 2009).

Our results indicate that the majority of practicing mathematics teachers do not regard the present contents of mathematics studies to be fully functional for future mathematics teachers. The practicing teachers suggested, for instance, separate courses for future mathematics teachers and mathematicians, and the possibility of taking school mathematics into account in the teaching of mathematics courses. These ideas were also proposed by some of the teacher educators. These findings are broadly in line with the well-known recommendations by other mathematics educators (e.g., Ball et al., 2008).

Both the practicing teachers and the teacher educators argued that course contents are purely general knowledge for future teachers if there were no explicit links with school mathematics, and hence these contents were not regarded as useful for future teachers. The practicing teachers argued that the links between university and school mathematics were difficult to perceive if the university course contents were set at too high a level compared with the usual school contents. The findings show that the pure mathematical contents, i.e., the common content knowledge (CCK) of the mathematics teacher education, should be carefully examined so that the most relevant mathematical contents for future mathematics teachers could be discovered. It is well known that a weak knowledge of mathematics on the part of teachers has a negative influence on teaching (McDiarmid, Ball & Anderson, 1989), but, on the other hand, a competence solely in mathematics is insufficient enough for good teaching (Hodgen, 2011). It seems that the relevance of mathematics courses depends on how explicitly the link between university and school mathematics is stressed in mathematics courses.

The results suggest that, in addition to pure mathematical contents, mathematics teacher education should include mathematical contents designed specifically for teachers. The practicing teachers argued that they needed mathematical knowledge and skills that were different from the skills and knowledge useful for mathematicians. This reflects the well-known ideas embodied in *pedagogical content knowledge* (Shulman, 1987), or *Specialized content knowledge*, *SCK* (Ball et al., 2008), i.e., an area of knowledge for teachers that also separates researchers from teachers. The practicing teachers and teacher educators suggested that the mathematical courses should be at least partly separate for future teachers and mathematicians. In practice, this would mean that more resources would be needed for mathematics teacher education, which might be a challenge.

Many of the teacher educators who participated in this study espoused the traditional view of development that emphasizes improving future teachers' subject matter knowledge (SMK) (Ball, 2003). Some educators viewed that good teaching requires knowledge of pure mathematics (CCK) and therefore they suggested that pure mathematical contents should be increased. On the other hand, many educators viewed that future mathematicians and future teachers need different kind of mathematical knowledge (SCK) and therefore they suggested that some of the present contents should be modified to be more suitable for teachers or new courses should be developed for teachers. Some educators viewed that forming integrated knowledge of pedagogy and mathematics (SCK) is one challenge for the mathematics teacher education, and therefore they suggested that the present courses should be re-scheduled or integrated.

On the other hand, a majority of the practicing teachers observed that there was a wider need for development than simply reforming the mathematical contents. The majority of practicing teachers demanded more courses concerned with teaching mathematics, students' learning difficulties in mathematics, and how to differentiate mathematics teaching. These knowledge domains can be identified as *Knowledge of content and teaching* (KCT) and *Knowledge of content and students* (KCS). The practicing teachers pointed out that they needed to alternate the knowledge and skills of teaching and learning in many classroom situations, and they seemed to consider that

the KCS and KCT knowledge types were interconnected especially in classroom *actions* (see Fernández, Figueiras, Deulofeu, et al., 2011; Ball et al., 2008). Many practicing teachers considered that they had learned pedagogical and mathematical issues in the course of their teacher education and that they had found teaching practice a very useful experience, but still they had difficulty in forming an integrated understanding of pedagogy and mathematics (see also Korthagen & Kessels, 1999; Sharp, 2004).

This linkage of theory and practice (Carlson, 1999; Tryggvason, 2009) seems to be a major challenge in mathematics teacher education, since it concerns not only the pedagogical and mathematical studies but also the teaching practice. Earlier research work has shown that solving the problem will not be simple. According to Verloop, Driel, and Meijer (2001), it is still difficult to foresee how teacher knowledge can be clarified clear for future teachers in their teaching practice. One of the problems appears to arise from the teacher educators' knowledge: not even experienced educators in the field of teaching practice have a clear grasp of the types of knowledge that teaching procedures involve, which makes it difficult to make the connection between theory and practice visible to student teachers (Verloop et al., 2001; Asikainen, Pehkonen & Hirvonen, 2013).

Filling the gap between theory and practice is a demanding task because teaching practice comprises only a small proportion of the teacher education studies as a whole. Hence, it is almost unrealistic to suggest that the gap could be fulfilled during the teaching practice. Our results suggest that the links between theory and practice should be made visible in all of the components of the teacher education so as to support future teacher development. Numerous suggestions have been made for the solution of this problem, e.g., by approaching it from practice to theory (Carlson, 1999), by developing the pedagogy of teacher education (Korthagen & Kessels, 1999), or by taking problem-solving into account in mathematics teacher education (Leikin & Levav-Waynberg, 2007). There is a possibility that contents, teaching methods, and the learning process may all be involved in the solution.

We have come to the realization that one of the key factors in reforming mathematical studies is the performance of a detailed analysis and comparison of curricula in university and school mathematics. In fact, it would seem obvious that the pure mathematical contents (CCK) should be the same as the topics in school mathematics or, at the very least, explicit links should exist between university and school mathematics. Another challenge is to design and develop special content knowledge (SCK) courses for future teachers. As yet, there is no general consensus about the knowledge and skills included in SCK (see e.g., Carrillo, Climent, Contreras & Muñoz-Catalán, 2013; Flores, Escudero & Carillo, 2013) but there should be no problem in designing new courses for future teachers, since there would be no harm caused if the subject matter and pedagogical contents are mixed. But as far as conceptualizing MKT is concerned, there is still work to be done to reach a consensus about this type of knowledge.

Although teacher education and teachers' knowledge are related (Darling-Hammond, Chung, & Frelow, 2002), more research into the challenges revealed by individual teacher education programs will be needed in order to construct a broader picture of this multifaceted phenomenon. Individual reports may act as an important part in this process by evaluating and improving mathematics teacher education before all of the universal challenges facing mathematics teacher education have been fully recognized. Although the present study has concerned only a single mathematics teacher education program, we would suggest that the following issues may prove to be more general challenges facing all mathematics teacher education programs:

- The connections between university mathematics and school mathematics are not self-evident for student teachers. Teachers need pure mathematical knowledge, e.g. Common Content Knowledge (Ball et al., 2008) and Subject Matter Knowledge (Shulman, 1986;

1987). However, student teachers may find that the mathematics studied at university level is too advanced and has no clearly visible connections to the mathematics taught in school.

- Specific mathematical knowledge is missed from teacher education, while the contents of mathematical courses focus too largely on pure mathematics. In addition to mathematical content knowledge, teachers also need specific mathematical knowledge, e.g. Specialized Content Knowledge (Ball et al., 2008) or School Mathematics (O'Meara, 2010), because they need to carry out a variety of different activities (e.g., producing teaching materials, formulating and marking exams) for which pure mathematical knowledge is insufficient.
- Teachers may have too few tools to be able to teach “good and poor” students at the same time. In the classroom teachers are simultaneously attempting to evaluate their students’ starting levels, to recognize their individual learning habits, and also to implement different teaching strategies that will match up to the pertaining situation. The knowledge required in these situations can be referred to as Pedagogical Content Knowledge (Shulman, 1986; 1987) or as both Knowledge of Content and Students and Knowledge of Content and Teaching (Ball et al., 2008).
- Courses in teacher education may be too theoretical (e.g., Carlson, 1999; Korthagen & Kessells, 1999). Student teachers may feel that mathematical and pedagogical courses and also teaching practice are too far removed from teachers’ actual work.

The results of this study encourage us in the development work of mathematics teacher education although the circumstances are still difficult at the starting point. The most demanding part has been and will be to evaluate what the personnel in mathematics teacher education teach and what kind of methods they use. We believe that assessment, feedback, and the teacher education personnel themselves and their cooperation are important components in the process of improving teacher education. It is common sense that there is always a possibility of improvement, and therefore the development must begin from critical thinking: what can we do better? With this article, we should like to encourage other researchers to evaluate and develop teacher education, and hence we would close with words that are too frequently dead and buried:

- *Without criticism, development dies* –

Acknowledgement

We would like to thank *The Finnish Cultural Foundation, North Karelia Regional Fund* for funding our research project *Evaluating and Improving Mathematics Teacher Education*.

References

- Asikainen, M. A., Pehkonen, E., & Hirvonen, P. E., (2013). Finnish Mentor Mathematics Teachers' Views of Teacher Knowledge Required for Teaching Mathematics. *Higher Education Studies*, 3(1), 79–91.
- Ball, D. L., & Forzani, F. (2009). The work of teaching and the challenge for teacher education. *Journal of Teacher Education*, 60(5), 497–511.
- Ball, D. L., Hill, H.C., & Bass, H. (2005). Knowing mathematics for teaching: Who knows mathematics well enough to teach third grade, and how can we decide? *American Educator*, 29(1), 14–17, 20–22, 43–46.
- Ball, D. L., Sleep, L., Boerst, T. A., & Bass, H. (2009). Combining the development of practice and the practice of development in teacher education. *The Elementary School Journal*, 109(5), 458–474.
- Ball, D. L., Thames, M.A. & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, 59(5), 389–407.
- Ball, D. L., & Bass, H. (2003). Toward a practice-based theory of mathematical knowledge for teaching. In E. Simmt & B. Davis (Eds.), *Proceedings of the 2002 Annual Meeting of the Canadian Mathematics Education Study Group*, (pp. 3–14). Edmonton, AB: CMESG/GCEDM.
- Carlson, H. L. (1999). From practice to theory: A social constructivist approach to teacher education. *Teachers and Teaching: theory and practice*, 5(2), 203–218.
- Carrillo, J., Climent, N., Contreras, L. C., & Muñoz-Catalán, M. C. (2013). Determining Specialised Knowledge For Mathematics Teaching. In B. Ubuz, C. Haser & M. A. Mariotti (Eds.), *Proceedings of the Eighteenth Congress of the European Society for Research in Mathematics Education* (pp. 2985–2994). Ankara, Turkey: European Society for Research in Mathematics Education.
- Darling-Hammond, L., Chung, R., & Frelow, F. (2002). Variation in Teacher Preparation How Well Do Different Pathways Prepare Teachers to Teach? *Journal of Teacher Education*, 53(4), 286–302.
- Delaney, S., Ball, D. L., Hill, H. C., Schilling, S. G., & Zopf, D. (2008). “Mathematical knowledge for teaching”: Adapting U.S. measures for use in Ireland. *Journal of Mathematics Teacher Education*, 11(3), 171–197.
- Fauskanger, J., Jakobsen, A., Mosvold, R., & Bjuland, R. (2012). Analysis of psychometric properties as part of an iterative adaptation process of MKT items for use in other countries. *ZDM*, 44(3), 387–399.
- Fennema, E. & Franke, L. M. (1992). Teachers' knowledge and its impact. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 147–164). New York, NY: Macmillan.
- Fernández, S., Figueiras, L., Deulofeu, J., & Martínez, M. (2011). Re-defining HCK to approach transition. In M. Pytlak, T. Rowland, & E. Swoboda (Eds.), *Proceedings of the Seventh Congress of the European Society for Research in Mathematics Education* (pp. 2640–2649). University of Rzeszów, Poland.
- Flores, E., Escudero, D., & Carrillo, J. (2013). A theoretical review of specialised content Knowledge. In B. Ubuz, C. Haser, & M.A. Mariotti (Eds.), *Proceedings of the Eighteenth*

Congress of the European Society for Research in Mathematics Education (pp. 3055–3064). Ankara, Turkey: European Society for Research in Mathematics Education.

- Hashweh, M. (2005) Teacher pedagogical constructions: a reconfiguration of pedagogical content knowledge. *Teachers and Teaching: Theory and Practice*, 11(3), 273–292
- 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., Ball, D. L., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education*, 39(4), 372–400.
- Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction*, 26(4), 430–511.
- Hill, H. C., & Lubienski, S.T. (2007). Teachers' mathematics knowledge for teaching and school context: A study of California teachers. *Educational Policy*, 21(5), 747–768.
- Hill, H. C. (2007). Mathematical knowledge of middle school teachers: Implications for the No Child Left Behind Policy initiative. *Educational Evaluation and Policy Analysis*, 29(2), 95–114.
- Hill, H. C., Ball, D. L., Blunk, M. Goffney, I. M., & Rowan, B. (2007). Validating the ecological assumption: The relationship of measure scores to classroom teaching and student learning. *Measurement: Interdisciplinary Research and Perspectives*, 5(2–3), 107–117.
- Hill, H. C., Ball, D. L., Sleep, L., & Lewis, J. M. (2007). Assessing Teachers' Mathematical Knowledge: What Knowledge Matters and What Evidence Counts? In F. Lester (Ed.), *Handbook for Research on Mathematics Education* (2nd ed.) (pp. 111–155). Charlotte, NC: Information Age Publishing.
- Hill, H. C., Dean, C., & Goffney, I. M. (2007). Assessing Elemental and Structural Validity: Data from Teachers, Non-teachers, and Mathematicians. *Measurement: Interdisciplinary Research and Perspectives*, 5(2–3), 81–92.
- Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, 42(2), 371–406.
- Hill, H. C., Schilling, S. G., & Ball, D. L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *Elementary School Journal*, 105, 11–30.
- Hodgen, J. (2011). Knowing and identity: a situated theory of mathematics knowledge in teaching. In T. Rowland & K. Ruthven (Eds.), *Mathematical knowledge in teaching* (pp. 27–42). Dordrecht: Springer.
- Korthagen, F., & Kessels, J. (1999). Linking theory to practice: Changing the pedagogy of teacher education. *Educational Researcher*, 28(4), 4–17.
- Leikin, R., & Levav-Waynberg, A. (2007). Exploring mathematics teacher knowledge to explain the gap between theory-based recommendations and school practice in the use of connecting tasks. *Educational Studies in Mathematics*, 66(3), 349–371.
- Markworth, K., Goodwin, T., & Glisson, K. (2009). The development of mathematical knowledge for teaching in the student teaching practicum. In D. S. Mewborn & H. S. Lee (Eds.), *Scholarly Practices and Inquiry in the Preparation of Mathematics Teachers* (pp. 67–83). San Diego, CA: Association of Mathematics Teacher Educators.

- McDiarmid, G. W., Ball, D. L., & Andersen, C. W. (1989). Why staying one chapter ahead doesn't really work: subject-specific pedagogy. *The national center for research on teacher education*. Issue paper 88-6.
- Meredith, A. (1995). Terry's learning: some limitations of Shulman's pedagogical content knowledge. *Cambridge Journal of Education*, 25(2), 175–187.
- O'Meara, N. (2010). *Improving mathematics teaching at second level through the design of a model of teacher knowledge and an intervention aimed at developing teachers' knowledge*. University of Limerick. Dissertation.
- Rowland, T., Turner, F., Thwaites, A., & Huckstep, P. (2009). *Developing Primary Mathematics Teaching: reflecting on practice with the Knowledge Quartet*. London: Sage.
- Schilling, S. G. (2007). The role of psychometric modeling in test validation: An application of multidimensional item response theory. *Measurement: Interdisciplinary Research and Perspectives*, 5(2–3), 93–106.
- Schilling, S. G., & Hill, H. C. (2007). Assessing measures of mathematical knowledge for teaching: A validity argument approach. *Measurement: Interdisciplinary Research and Perspectives*, 5(2–3), 70–80.
- Schilling, S. G., Blunk, M., & Hill, H. C. (2007). Test validation and the MKT measures: Generalizations and conclusions. *Measurement: Interdisciplinary Research and Perspectives*, 5(2–3), 118–128.
- Sharp, J. (2004). Spherical Geometry as a Professional Development Context for K-12 Mathematics Teachers. In T. Watanabe & D. R. Thompson (Eds.), *The Work of Mathematics Teacher Educators: Exchanging Ideas for Effective Practice* (pp. 103–118). Association of Mathematics Teacher Educators.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14.
- Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57, 1–22.
- Sleep, L. (2009). *Teaching to the mathematical point: Knowing and using mathematics in teaching*. University of Michigan. Dissertation.
- Stylianides, A. J., & Ball, D. L. (2008). Understanding and describing mathematical knowledge for teaching: Knowledge about proof for engaging students in the activity of proving. *Journal of Mathematics Teacher Education*, 11(4), 307–332.
- Thames, M. H., & Ball, D. L. (2010). What mathematical knowledge does teaching require? Knowing mathematics in and for teaching. *Teaching Children Mathematics*, 17(4), 220–225.
- Tryggvason, M-T. (2009). Why is Finnish teacher education successful? Some goals Finnish teacher educators have for their teaching. *European Journal of Teacher Education*, 32(4), 369–382.
- Verloop, N., Van Driel, J., & Meijer, P. (2001). Teacher knowledge and the knowledge base of teaching. *International Journal of Educational Research*, 35(5), 441–461.

Use of Mathematical Tasks of Teaching and the Corresponding LMT Measures in the Malawi Context

Mercy Kazima
University of Malawi, Malawi

Arne Jakobsen
University of Stavanger, Norway

Dun Nkhoma Kasoka
University of Malawi, Malawi

Abstract: We discuss the adaptation and piloting of the previously developed U.S.-specific measures of mathematical knowledge for teaching to the Malawi context. The purpose is to produce measures that can be used to evaluate changes in mathematical knowledge for teaching gained through primary teacher education, thus informing teacher educators on the most effective evidence-based practices. By interviewing 14 teachers, we first examine whether the 16 recurrent mathematical tasks of teaching tasks identified in the U.S. are applicable to the Malawi context. This is followed by the discussion of the adaptability of the U.S. developed number concept and operations LMT measures. Next, we report on the item psychometric properties estimated from a pilot study in which 351 preservice primary school teachers participated at the end of their coursework. Our findings suggest that all the 16 tasks of teaching mathematics are applicable to the Malawi context, albeit to varying degrees, and should be complemented by additional tasks suggested by the Malawi teachers. For the LMT measures, we found that the majority of the LMT items psychometrically function well in the Malawi context and that item difficulty estimated in Malawi was strongly correlated with that reported in the U.S. We thus argue that there is some generality to the mathematics teaching tasks across the two contexts, as well as some specificity to Malawi, and that the adapted LMT measures can be used in a Malawi context.

Keywords: Teacher knowledge, Mathematical knowledge for teaching, Primary teacher education, Malawi.

Introduction

Teacher knowledge is important for both teaching and learning. Since Shulman (1986) introduced the concept of pedagogical content knowledge, his ideas have triggered widespread interest among researchers and practitioners alike. In addition, different conceptualizations of teacher knowledge have emerged, such as Knowledge Quartet (Rowland, Huckstep, & Thwaites, 2005), Knowledge for Teaching (Davis & Simmt, 2006), and Mathematical Knowledge for Teaching (Ball, Thames, & Phelps, 2008). Despite different views on categorizations, researchers seem to agree that teacher knowledge plays a key role in student learning (e.g., An, Kulm, & Wu, 2004; Rowan, Correnti, & Miller, 2002; Wright, Horn, & Sanders, 1997), and much progress has been made in understanding the professional mathematical knowledge that teachers need in order to perform the recurrent tasks of teaching mathematics. There is also an extensive body of knowledge on how this knowledge is acquired, how it can be measured, and how it relates to teaching and student learning (e.g.,

Ball et al., 2008; Hill, Rowan, & Ball, 2005; Rowland et al., 2005). Nonetheless, a greater understanding of how this knowledge might differ as context changes is necessary.

Different approaches to measuring teacher knowledge are presently in use, one of which is based on analyzing school mathematics curriculum and developing measures that would test knowledge and the teaching of that curriculum. The main drawback of this approach is that it requires many, often inaccurate, assumptions about outcomes of teaching that is aligned with the curriculum. Ball and colleagues (2008) developed an alternative practice-based approach in the United States. According to the authors, their aim was to “unearth the ways in which mathematics is involved in contending with the regular day-to-day, moment-to-moment demands of teaching” (p. 395). Using this approach enabled these researchers to define the theory of mathematical knowledge for teaching (MKT) (Ball & Bass, 2003). In particular, they identified a list of 16 mathematical teaching tasks that are part of the work teachers routinely do (Ball et al., 2008). In connection to this work, measures of mathematical knowledge for teaching were developed as a part of the Learning Mathematics for Teaching (LMT) project at the University of Michigan. The LMT project developed items in three content areas, namely (i) number concepts and operations (NCOP); (ii) geometry; and (iii) patterns, functions, and algebra (a sample of released items can be found in Ball and Hill (2008)).

The LMT measures¹ have been adapted for use in different contexts outside the U.S., for example, in Ireland (Delaney, 2012), Norway (Fauskanger, Jakobsen, Mosvold, & Bjuland, 2012), Indonesia (Ng, 2012), South Korea (Kwon, Thames, & Pang, 2012), and Ghana (Cole, 2012). In South African context, Adler and Patahuddin (2012) researched mathematical knowledge for teaching, reporting that the LMT items “have much potential in provoking teachers’ talk and their mathematical reasoning in relation to practice-based scenarios; and exploring with teachers a range of connected knowledge related to the teaching of a particular concept or topic is most important resource for teachers” (p. 17).

It is important to note that the mathematical knowledge for teaching theory and the associated LMT measures were developed from classroom observation in the U.S. and were not intended for use in other cultures. Thus, we were interested in exploring whether the 16 tasks of teaching identified by Ball et al. (2008) also are applicable in a Malawi context. By applicable we mean if teachers in Malawi would do the same or similar tasks as part of their work of teaching mathematics in schools. Our interest in mathematical knowledge for teaching, with associated tasks of teaching and measures, developed from our work in teacher education in Malawi. With the overarching project goal of improving quality and capacity of mathematics teacher education in Malawi, we are interested to learn more about the development of preservice teachers’ mathematical knowledge for teaching through their teacher education. Knowing what works and what does not in teacher education in Malawi is the first step in the process of improving the quality of teacher education. In our view, measures of mathematical knowledge for teaching can help answer these questions. In this larger study, we plan to measure preservice teacher mathematical knowledge for teaching before and after coursework, using a pre- and post-test. This can help us determine knowledge growth in preservice teachers, and consequently inform our practice. Since the LMT measures were developed from studying work and identifying tasks of teaching in the U.S. context, our first objective was to ascertain whether these tasks are applicable to the Malawi context.

¹ The measures (or instruments) developed as part of the Learning Mathematics for Teaching Project at the University of Michigan are also sometimes called LMT items, MKT items, or MKT measures. In this paper, will use the terms LMT measures/instruments.

These findings elucidated the suitability of adapting the LMT measures for use in Malawi. Our next goal was to outline the method of adaptation that the measures underwent. This paper is therefore guided by the following research questions:

1. Are the tasks of teaching mathematics identified in the U.S. applicable in Malawi?
2. What can we learn about the adaptability of LMT measures to the Malawi context from psychometric properties estimated in a pilot study?

Other researchers have raised similar questions when adapting the LMT measures to their contexts (for example, Cole, 2012; Delaney, 2012; Fauskanger et al., 2012; Ng, 2012). We focus specifically on the Malawi context, acknowledging the cultural differences between Malawi and other cultures. Furthermore, in their work, Hoover, Mosvold, and Fauskanger (2014) called for “increased efforts to identify professionally defensible mathematical tasks of teaching that can serve as a common foundation for conceptualizing and measuring mathematical knowledge for teaching internationally” (p. 7). Our study is an attempt to respond to the call. By adapting the measures and piloting a set of 88 items, our study will also make an argument for use of the LMT measures in a Malawi context.

Adapting LMT Measures to Different Cultural Contexts

Many researchers agree that teacher knowledge of mathematics is influential in shaping teaching practices that in turn affect student achievement (e.g., Hill, Ball, & Schilling, 2008). One of the reasons the work of Ball and colleagues has received extensive attention is that the authors based their findings on a thorough study of actual classroom teaching. Because their research was conducted in the U.S. context, it is debatable whether the findings are generalizable. For example, Andrew (2011) argued that “teacher’s mathematical knowledge, as manifested in their observable behavior, is a cultural construction” (p. 100). While we raise similar questions about generalizing the theory of mathematical knowledge for teaching to the Malawi context, we also draw upon findings of similar studies in this field. We agree with the view that the practice of teaching is a cultural activity, making teaching of mathematics culturally specific, even though the content is general (Delaney, 2012).

According to the results reported by Cole (2012), the LMT measures can be used in the Ghanaian context after careful adaptation. However, so far Ghana is the only African country where the LMT measures have been used. Moreover, Cole and other researchers have demonstrated that adaptation of LMT measures is not easy and requires a process that considers all differences between the contexts, including culture, language, and teaching practices (Delaney, 2012; Fauskanger et al., 2012). In our study, we observed that the differences were manageable and hence proceeded with the adaptation.

Malda et al. (2008) emphasized on fairness of adapted measures, noting that “it is unfair to assess intelligence of children from Africa with a test that has been validated in a Western culture . . . , with a population of children exposed to very different educational and material environments at home and school” (p. 452). Similarly, it could be argued that it is not fair to assess Malawian teachers’ knowledge using measures developed for the vastly different U.S. context. As Hoover et al. (2014) pointed out, such an argument focuses on differences in the practice of teaching in different cultures. The authors further argued that, to evaluate such arguments, we need to consider “the underlying concept of *work of teaching* and *tasks of teaching* used in MKT assessment items and to ask whether, or to what extent, such concepts, as defined therein, are meaningful across cultural contexts” (p. 8) [emphasis in original]. In this study, we consider whether the items we adapted are meaningful in the Malawi context.

Malawi Context

Malawi, formerly known as Nyasaland, was a British protectorate that gained its independence from Britain in 1964. English remains the official language and is also the school language from fifth year of primary school onwards. Malawi's school system comprises of eight years of primary and four years of secondary school. Primary school education is free and easily accessible, neither of which is the case for secondary school education (Kazima & Mussa, 2011). Currently recommended age for enrollment into the first grade of primary school is six years. Thus, primary school students are aged 6 to 13 years old, whereas secondary school serves 14 to 17-year-olds. However, in reality, many classrooms are attended by children of various ages because some children commence education when they are older than six and some repeat classes. At the end of secondary school, students take Malawi Schools Certificate of Education (MSCE) national examinations, equivalent to the Ordinary level (General Certificate of Education). Tertiary education, including teacher education, requires passing the MSCE examinations. Teacher education for primary schools is offered by teacher education colleges, most of which are government owned, while some are private institutions. All teacher colleges follow one curriculum, referred to as the Initial Primary Teacher Education program. It is a two-year full time program, comprising of college coursework in the first year and teaching practice, which is the focus of the second year (Malawi Institute of Education, 2010). Students that successfully complete the program are awarded Primary Teacher's Certificate. There is no subject specialization for primary teachers; all preservice teachers learn all subject areas, including mathematics, and they are expected to teach all subjects in primary schools (Malawi Institute of Education, 2010). As a part of this study, we piloted the adaptability of LMT measures on preservice teachers at the end of their coursework, as a part of the first year of the teacher education program.

Methods

Mathematical Tasks of Teaching

In order to answer our first research question about the applicability of U.S.-specific tasks of teaching in the Malawi context, we asked 14 experienced teachers to respond to a questionnaire and followed this with a group discussion. All study participants were practicing teachers from various primary schools in Malawi and were drawn from an in-service upgrading course at the University of Malawi. The questionnaire consisted of two parts, whereby Part A listed all the 16 core tasks of teaching mathematics (Ball et al., 2008). The respondents were asked to indicate which ones were applicable to them as mathematics teachers and were required to elaborate on their responses. Verbal instructions were given to the teachers prior to completing the questionnaire. In particular, they were informed that they should draw from their experience and consider whether each of the tasks is something they do as a part of their work of teaching mathematics. In Part B of the questionnaire, the teachers were instructed to note tasks of teaching mathematics that they perform that are not included on the list. All study participants completed the questionnaire at the same time, in one room. Subsequently, eight teachers took part in a group discussion.

LMT Measures

We answered the second research question in two phases—Phase I, which was the actual adaptation of instruments, and Phase II, pilot testing of the adapted instrument.

Phase I: Adapting measures. Phase I was performed in three stages, the first of which consisted of selecting the most appropriate instruments from those available. We limited our study to measures of number concepts and operations (NCOP) content area. The

LMT project provided us with ten forms,² with items from the content area number concept and operations, and corresponding questions related to content knowledge (CK) and knowledge of content and students (KCS) (see, for example, Ball et al., 2008). Upon closer inspection, we noted some repetition and modification of stems and items from the 2001 to the 2004 version. After aligning items in each form to Malawi's mathematics curriculum for Initial Primary Teacher Education, we observed that Form A from the 2001 instrument (NCOP-CK_2001A) had the closest and the most comprehensive range of items covering the curriculum. Therefore, we selected this form and its corresponding Form B (NCOP-CK_2001B) as a starting point for adaption. There were six common items between the two forms and because we wanted to pilot as many items as possible, we replaced the common items in Form B with similar items from the remaining forms. We also added new stems to both forms to cover the concepts of division, multiples, and factors, which were not sufficiently covered by the original forms. These stems were also taken from some of the remaining forms. In total, our new Form A had 25 stems and 46 items, while Form B had 24 stems with 42 items, making a set of 88 items. The original 2001 Form A has 13 stems with 26 items, while Form B has 15 stems with 24 items. Having two forms at this stage allowed us to pilot as many items as possible. We considered the alternative of having all items in one form, but decided that it was not appropriate, since it would make the form too long. Another argument for having two forms at this stage was that the project would eventually need to group all items into two comparable forms that could be used in pre- and post-tests.

The second stage of adaptation involved contextualizing the U.S.-based instruments, both stems and items, to the Malawi context. This was done by changing some words, phrases, and names of places, people, and objects to what we think would be familiar in Malawi context, as shown in Table 1.

Table 1. *Examples of Changes Made from U.S. to Malawi Context.*

Category of change	U.S.	Malawi
General context: • Names of people (42 instances)	Ms. Jamison	Mrs. Banda
	Chad	Chisomo
• Names of objects (21 instances)	Pizza	Bread
	Field	Farm
School context: (11 instances)	Scoring/Reviewing	Marking
	Rules of thumb	Simple rules
	Student papers	Student notebooks
	State assessment	MANEB examinations
	Quiz	Test
	Mini-lessons for students focused on particular difficulties	Revision lessons
Item Numbering	Numbering not continuous	Continuous numbering

² These are three 2001 NCOP-CK forms (A, B, and C), three 2001 NCOP-KCS forms (A, B, and C), two 2002 NCOP-CK forms, and two 2004 NCOP-CK forms.

The third and final stage required checking and modifying the mathematical content of each stem and the corresponding items in order to ensure that they reflect the Malawi curriculum. The modifications made addressed “changes related to school cultural context” and “changes related to mathematical substance” (Delaney et al., 2008, p. 182). After initial changes were made to both forms, we sought input from four experienced primary school teachers via semi-structured individual interviews. Their feedback helped us modify the items further, thus ensuring relevance of the content, wording, representations, and notations (Kasoka, Kazima, & Jakobsen, 2016).

Phase II: Pilot testing. The two adapted (and extended) forms (Form A and B) were subjected to a pilot test by 351 preservice primary school teachers. The two forms were administered to all preservice teachers at one Teacher Education College at the end of their first-year coursework. They were all informed of the objectives of the study and the instructions were carefully explained to them. The two forms were randomly distributed to the students, whereby 212 students answered Form A while 139 answered Form B. Ideally, we wanted the two groups to be equal in size, but due to incorrectly estimating the total number of students at the teacher college, this was not achieved. We had no control over the distribution of the two forms within the group because the forms were mixed in advance for randomness. While the preservice teachers were allowed as much time as they needed to complete the forms, the majority took 1–2 hours. After the test, fifteen randomly selected students were asked to comment on the test and the specific items. Their input allowed us to assess the suitability of the items for testing teacher knowledge in Malawi.

Results and Findings

Mathematical Tasks of Teaching

Table 2 shows the results of Part A of the questionnaire answered by the 14 teachers. The results reported pertain to the frequencies, the number of teachers that indicated that the task was applicable (yes), or not applicable (no), or did not respond to the question (non-response).

As can be seen from Table 2, there was no task that none of the teachers identified with, and all 16 tasks were considered applicable to Malawi school context by at least five of the teachers. In other words, the teachers viewed these tasks as something they do as a part of their work of teaching mathematics. This finding seems to suggest that U.S. and Malawi contexts share some similarities. However, there were variations in the frequencies, as some tasks were considered applicable by a greater number of teachers than others were. Thus, in subsequent analysis, we consider these tasks as most applicable to the Malawi context. We also identified three tasks that all teachers felt were applicable to Malawi, namely “presenting mathematical ideas,” “finding an example to make a specific mathematical point,” and “appraising and adapting the mathematical content of textbooks.” All these tasks involve typical traditional classroom practices of mathematics teachers—explaining mathematics content or procedure, illustrating to students using examples and often from textbooks, and asking students to practice exercises from textbooks. Hence, it is not surprising that all the teachers identified these as aspects of the work they do when teaching mathematics, thus making these tasks most applicable to the Malawi school context.

Table 2. *Frequencies of Responses to Applicability of Tasks to the Malawi School Context.*

Task Description (from Ball et al., 2008)	Is task applicable to Malawi school context?		
	Yes (%)	No (%)	Non-response (%)
Presenting mathematical ideas	14 (100)	0 (0)	0 (0)
Finding an example to make a specific mathematical point	14 (100)	0 (0)	0 (0)
Appraising and adapting the mathematical content of textbooks	14 (100)	0 (0)	0 (0)
Connecting a topic being taught to topics from prior or future years	13 (93)	0 (0)	1 (7)
Recognizing what is involved in using a particular representation	12 (86)	1 (7)	1 (7)
Modifying tasks to be either easier or harder	12 (86)	1 (7)	1 (7)
Responding to students' "why" questions	11 (79)	3 (21)	0 (0)
Linking representations to underlying ideas and to other representations	11 (79)	2 (14)	1 (7)
Giving or evaluating mathematical explanations	10 (71)	2 (14)	2 (14)
Choosing and developing useable definitions	10 (71)	2 (14)	2 (14)
Asking productive mathematical questions	10 (71)	2 (14)	2 (14)
Selecting representations for particular purposes	10 (71)	2 (14)	2 (14)
Evaluating the plausibility of students' claims (often quickly)	9 (64)	2 (14)	3 (21)
Using mathematical notation and language and critiquing its use	8 (57)	3 (21)	3 (21)
Inspecting equivalences	6 (43)	5 (36)	3 (21)
Explaining mathematical goals and purposes to parents	5 (36)	8 (57)	1 (7)

In sum, 12 of the 16 tasks were identified as applicable by more than 70% of the teachers. Of the remaining four tasks with lower frequencies, only two—"inspecting equivalences" and "explaining mathematical goals and purposes to parents"—were identified as applicable to the Malawi context by less than half of the teachers. Those that deemed the task applicable explained that they discuss teaching mathematics (and other subjects) with parents during gatherings, such as Parents and Teachers Association meetings. During the subsequent group discussion, it was clear that some schools do indeed inform parents about the teaching of mathematics.

It is also important to note that not all teachers responded to all questions and some failed to elaborate on their responses. This issue was raised in the group discussion, where the participants revealed that, if they were unsure about some of the tasks, they felt that it was

appropriate not to respond to the question.

Tasks Teachers Did Not Find Applicable

We also examined the tasks that some teachers did not view as something they do. Looking at Table 2, it is evident that, for 12 of the 16 tasks, at least one teacher responded that he/she does not see it as applicable to the Malawi context. We have presented some of these tasks in Table 3 in descending order of frequencies. We have also included examples of the reasons that teachers gave. It is important to note that, following the group discussion with eight teachers, no new information arose. Nevertheless, the discussion helped explain some of the written responses.

Table 3. *Examples of Teachers' Reasons for Tasks Not Being Applicable to Malawi Contexts.*

Mathematical task of teaching	Freq. for not applicable	Reasons
Explaining mathematical goals and purposes to parents	8 (57%)	Parents rarely know what their children learn in school. Majority of teachers do not have time to deal with parents. Parents are not involved in class work. Learners' guardians do not ask about mathematics. Teachers do not discuss mathematics with parents.
Inspecting equivalences	5 (36%)	Teachers' guide matches content to find links Associating concepts is not done by teachers. We do not do this in primary school.
Responding to students' "why" questions	3 (21%)	Teachers just give explanations. Learners ask how to solve problems. Teachers do not do this.
Using mathematical notation and language and critiquing its use	3 (21%)	Teachers do not critique teachers' guide. Notation given in textbooks. Guided by teachers' guide.
Choosing and developing useable definitions	2 (14%)	Definitions are given in textbooks. Teachers guide has definitions.
Selecting representations for particular purposes	2 (14%)	Teachers' guide has representations. Use text books.
Modifying tasks to be either easier or harder	1 (7%)	Textbook includes both easy and hard problems.

More than half of the teachers indicated that the task of "explaining mathematical goals and purposes to parents" is not applicable to the Malawi school context, as this is not something they do as a part of their work. The reasons given suggest that the teachers do not consider this as a mathematical task of teaching because, according to the teachers, parents are not involved in what goes on in the mathematics classroom; thus, teachers do not talk to parents about mathematics. During the group discussion, some teachers said that they have

not seen this practice in any school they had worked in. They explained that they have meetings with parents where they discuss students' welfare, but do not address any academic issues. From our experience of the Malawi context, some schools would not involve parents into their students' academic progress due to the low literacy levels of parents and guardians, especially in rural areas. Another reason might be teachers' lack of subject matter knowledge or confidence in their own knowledge of mathematics. This was not mentioned by the teachers that took part in this study, but seems plausible that, if teachers are not confident of their knowledge or are aware of their limitations, they might not be willing to discuss their work with parents. Interestingly, one teacher stated, "learners' guardians do not ask about mathematics," which seems to suggest that, if the parents asked about goals and purposes of mathematics, the teacher would explain them. Another teacher noted, "majority of teachers do not have time to deal with parents," suggesting that lack of time is the only reason for this omission. Nevertheless, it is evident that some teachers do not perceive parent involvement as a task related to teaching mathematics in Malawi primary schools.

The mathematical task of "inspecting equivalences" was found not applicable by five teachers, two of whom did not give any reasons. The remaining three explained that this is done for the teachers but not by the teachers themselves. One teacher clearly stated that this is done by the teachers' guides, which are books for teachers that accompany textbooks for each grade.

Most of the reasons given pertain to teachers' guides and textbooks, which are perceived as sufficient, thus absolving the teachers from the responsibility for checking the equivalencies, as this is done by the authors of these curriculum materials. For example, the tasks of "using mathematical notation and language and critiquing its use," "choosing and developing useable definitions," and "selecting representations for particular purposes" were not perceived by the teachers as something they should do, as they are guided by what is written in textbooks and in teachers' guides. It appears that the teachers follow the textbook and teachers' guide diligently without questioning their applicability. This is common among Malawi teachers, who tend to take teachers' guide and textbooks as prescriptions of what and how to teach and not as suggestions. It might be important to note that the Malawi Institute of Education—a government institution—is the sole provider of textbooks and accompanying teachers' guides for primary schools in Malawi and these are made available to the teachers. Hence, for most of the teachers, these are the only resources they have and use. There are other possible reasons for not identifying with the tasks, although these are not mentioned by the teachers. For example, limited knowledge of the mathematical content by the teachers would compromise their ability to see equivalences, develop useable definitions, select representations, or critique use of mathematical notations. All these require deep understanding of the mathematics involved and might explain the heavy reliance on teachers' guides.

Similar to reasons for finding tasks applicable to the Malawi context, some teachers did not give the reason for selecting "no" as their response. It is likely that these teachers did not clearly understand the task or their own perception of it, and thus felt that it would not be appropriate to elaborate on their response.

Other Mathematical Teaching Tasks

Table 4 shows results pertaining to Part B of questionnaire, where teachers were asked to indicate tasks of teaching mathematics in Malawi schools that are not included in the original list of 16.

Table 4. *Frequencies of Responses to Other Mathematical Teaching Tasks in the Malawi Context.*

Other mathematical tasks of teaching mentioned	Frequency (%)
Making teaching and learning aids (TALULAR)	14 (100)
Assessing learners (preparing tests, marking scripts)	10 (71)
Using different teaching methods	8 (57)
Ordering content/sequencing topics	6 (43)
Relating mathematics to context and other subjects	4 (29)

As can be seen from the results presented in Table 4, five additional tasks of teaching were mentioned by at least four of the teachers. All participating teachers mentioned “making teaching and learning aids” and they referred to this as TALULAR, which stands for Teaching And Learning Using Locally Available Resources. TALULAR is a common concept in Malawi schools and is emphasized in teacher education. Thus, teachers take it as part of their job to make their own resources for teaching (Malawi Institute of Education, 2010). We acknowledge that use of teaching aids per se is not part of Ball et al.’s (2008) conceptualization of mathematical tasks of teaching. However, we contend that, in the Malawi context, where the teachers are expected to make the teaching aids from their local resources, this involves some mathematical reasoning and is hence integral part of their mathematical tasks of teaching.

The next common additional task was “assessing learners,” most likely because assessment of learners is very important in Malawi schools, where learners take school tests at the end of each term, as well as during the term. Furthermore, learners’ progression from one grade to the next is dependent on passing such tests. It is therefore not surprising that Malawi teachers consider “assessing learners” as one of distinct mathematical tasks of teaching. Indeed, the reasons teachers gave for explicitly noting this task was that it is their job to prepare tests, evaluate learners’ work, and give feedback to their students. However, this is not a distinct task because the work involved in assessing students is reflected in some of the tasks in the original list of 16, such as “evaluating the plausibility of students’ claims” and “asking productive mathematical questions.”

More than half of the teachers mentioned “using different teaching methods” as another distinct mathematical task of teaching, as all respondents felt that it was linked to using learner-centered approaches in teaching. This reflects the current curriculum and initiatives in Malawi schools, which emphasize learner-centered teaching, as opposed to the traditional teacher-centered teaching methods (Malawi Institute of Education, 2014). We also acknowledge here that teaching methods do not fit with the conceptualization of mathematical tasks for teaching by Ball et al. (2008).

The last two suggested tasks “relating mathematics to context and other subjects” and “ordering content/sequencing topics” are partly covered by the task of “connecting a topic being taught to topics from prior or future years.” While during the group discussion, teachers acknowledged this, they nonetheless felt that their suggestions extended beyond what is covered in the 16 tasks. For example, comparing “connecting a topic being taught to topics from prior or future years” with “relating mathematics to context and other subjects,” one teacher said:

. . . that task is about scope and sequence, what we teach before and after . . . it is

talking about scope and sequence . . . this we are saying linking to everyday life and other learning areas, science, agriculture . . . it is not same thing

The fact that the teachers were suggesting other tasks that are partly covered in the 16 tasks emphasizes our finding from Part A that, in general, the teachers find the tasks originally developed in the U.S. applicable to the Malawi context. The parts that are not covered seem to be additional tasks for Malawi and therefore worth paying attention to. We acknowledge that Ball et al. (2008) do not claim that the list of 16 is exhaustive, and our goal is not to show that the Malawi context has more tasks. Rather, our aim is to illustrate that there seems to be notable similarities between the two contexts, while acknowledging some specificities of the Malawi context.

Discussion on Mathematical Tasks of Teaching

These findings are drawn from teachers' responses to the questionnaire and the information gained during the group discussion, rather than from observations of actual teaching of mathematics in Malawi schools. Some might dispute this approach because it is based on teachers' views and not their practice. While we acknowledge the limitations of our approach, we argue that experienced teachers would be able to consider each of the tasks presented to them and draw from their experience to determine whether the task is something they do as part of their work of teaching mathematics. We emphasize that the goal of this work was to gain a better understanding of what teachers perceive as tasks of teaching in the Malawi context. On one hand, the fact that all 16 tasks were recognized by at least a third of the teachers informs us that the recurrent tasks of teaching are generally applicable to both the U.S. and the Malawi context. On the other hand, the fact that some tasks were not recognized by a relatively large number of teachers suggests that there is some specificity to the Malawi context. Thus, our findings help elucidate some differences between the two contexts, which support the claim that teaching is culturally specific (Delaney, 2012). The additional tasks of teaching described by teachers, especially the making of teaching aids (which was mentioned by all teachers), emphasize further the cultural specificity of teaching mathematics in the Malawi context. Thus, while we acknowledge that Ball et al. (2008) do not claim that the list of 16 is exhaustive, we also highlight the fact that some of the additional tasks the Malawi teachers suggested are covered in the original list of 16.

Piloting Adapted Measures

As explained earlier, the adapted forms were piloted on preservice teachers, whose responses were analyzed via the BILOG-MG software version 3.0 (Zimowski, Muraki, Islevy, & Bock, 2003) using the 2-parameter logistic (2PL) IRT model. As suggested by Hambleton, Swaminathan, and Rogers (1991), a model-data fit was investigated using multiple methodologies, starting from the assumptions of the IRT model, including unidimensionality. Since the two forms were answered by two equivalent groups (drawn from the same sample), we used equivalent groups equating with no overlapping items, placing the items on the two forms on the same scale. The 2PL IRT model produces two parameters for each item—parameter a that defines item discrimination, and parameter b , often referred to as item difficulty, as it shows the location where a respondent with ability of b will have a 50% chance of answering the item correctly (Edwards, 2009). Item characteristics on this scale typically range from -3 to $+3$, with items associated with higher values being more difficult than those with lower values. We used the values of these parameters to investigate problematic items in the Malawian context, and to identify items that are candidates for inclusion in the two forms that can be used for measuring teacher knowledge. We also studied point biserial correlation to identify problematic items (Fauskanger et al., 2012; Ng, 2012).

All the preservice teachers were also given an IRT score, placing them along the ability interval with 0 as the mean ability level and 1.0 as the standard deviation. It is, however, important to note that the preservice teachers' scores are not the focus of this paper, even though the psychometric properties of the items are.

The results yielded by the IRT analysis revealed an IRT reliability of 0.874, whereby maximum information was obtained at 1.25 standard deviation above the mean score. This indicated that this set of items is optimal for assessing more knowledgeable preservice teachers. BILOG flagged two items that had problematic point biserial (less than -0.15) and these were omitted from further calculation (Item 11b and 24d, both from Form B). In the next phase of data analysis, we examined the relative item difficulty distribution reported in the U.S. and estimated in the Malawi context. Even though the two scales are not directly comparable, one can expect—if the items function similarly in the Malawi as in the U.S. context—that there are some similarities among the relative ordering of item difficulties when looking at the full set of items (Fauskanger et al., 2012). More specifically, we posit that, if items are ordered from easy (low value of the b parameter/difficulty) to hard (more difficult, high value of the b parameter), one should expect—if the psychometrical properties of these items are similar across the two populations—that the order should be maintained (Fauskanger et al., 2012). The strong correlation pertaining to the relative item difficulty in the U.S. and Malawi context indicates that the relative order of the item difficulties is strongly maintained. This finding suggests that the items seem to function in a similar manner in the two populations.

For this set of items (Form A and Form B), the item difficulty found in the U.S. ranged from -4.281 to 3.961, with the average item difficulty of -0.535. In Malawi, the relative item difficulty ranged from -5.56 to 4.98, with the average difficulty of 1.119. A scatterplot of relative item difficulty found in the U.S. and in Malawi is depicted in Figure 1.

Figure 1. A scatterplot of the relative item difficulty found in the Malawi context, plotted against item difficulty found in the U.S.

We found a strong correlation between the item difficulty reported in the U.S. and that

estimated in the Malawi context (Pearson correlation $r = .738$, $p < .0005$). The average item difficulty for this set of items was -0.535 for the U.S., while 1.119 was estimated in the Malawi context.

Discussion of the Psychometric Results

The primary purpose of the psychometric analysis was to ensure that the adapted items function well psychometrically in the Malawi context, and to finalize two forms with items—including anchoring items—that can be used for further studies in Malawi. For the set of items used in this pilot, it seems clear from the results that some of the more difficult items should be taken out of the set. Firstly, the point of maximum information (0.874) and the average item difficulty (1.119) indicated that the set of items was skewed in favor of the more knowledgeable respondents, implying that, overall, there were too many difficult items in the set. We were ideally seeking to compose two forms, both with approximately 30 items (including five repeating items as anchoring items) in a pre- and post-test study. Because of this aim, we decided to remove some items that appeared too difficult for Malawi preservice teachers. The main objective was to attain item distribution on each form with majority of items around mean ability of zero. As a result of these considerations, Item 2b, 3, 5b, 6, 7, 13, 18, 21e, and 24 from Form A, and Item 3, 8, 10, 11b, 14b, 16a, 17, 18, 20, 21, 22, and 24d from Form B were removed. We then added five items from our Form B to Form A, and one item from Form A to Form B. After this Form A had 38 items and Form B 35 items including six anchoring items on both forms. The resulting revised set of items had an average item difficulty of 0.854 in Malawi, versus -0.770 obtained in the U.S. The final item distribution of the new forms is given in Table 5 below, which also provides the average difficulty level when the items are distributed to the two forms.

Table 5. *Distribution of Item Difficulty for the Malawi Context.*

Difficulty level	Number of items		
	Form A	Form B	Common items
$b < -3$	2	1	
$-3 \leq b < -2$	-	1	
$-2 \leq b < -1$	4	4	1
$-1 \leq b < 0$	7	7	1
$0 < b \leq 1$	9	9	
$1 < b \leq 2$	10	8	3
$2 < b \leq 3$	5	4	1
$b > 3$	1	1	
Total items	38	35	6
Total stems	20	20	6
Average Malawi	0.823	0.425	
Average U.S.	-0.83	-0.929	

Conclusion

In conclusion, we revisit our research questions and briefly explain how we have answered them in this work based on the findings we obtained. With respect to the first question—*Are the tasks of teaching mathematics identified in the U.S. applicable in*

Malawi?—our findings show that the 16 tasks are applicable in the Malawi school context. However, some of the tasks are more commonly recognized by teachers as applicable to the Malawi context, while other tasks are found less relevant. The Malawi school context is very different from the U.S. context and it is interesting that the work of teaching seems generally similar. It is also intriguing to observe that, while there seems to be generality between the two contexts, there is also some specificity to the Malawi setting. In Malawi, making teaching aids from local resources is seen as integral part of teaching mathematics. It is important to note that this study does not draw any inferences about the way the tasks are carried out. It also does not extend to the assessment of their implementation by teachers in the Malawi classroom. Thus, while answering the call for efforts to identify mathematical tasks of teaching as a basis for conceptualizing MKT (Hoover et al., 2014), the study does not provide empirical evidence. However, it does offer a valuable foundation for further studies in this field, where empirical evidence can be gathered in order to contribute towards building a shared understanding of “common tasks of teaching mathematics” and consequently development of “internationally shared measures of MKT” (Hoover et al., 2014, p. 9).

For the second research question—*What can we learn about the adaptability of LMT measures to the Malawi context from psychometric properties estimated in a pilot study?*—we found, after piloting the set of 88 adapted items, that majority of items psychometrically function well in the Malawi context. The item difficulty reported in the U.S. was strongly correlated with that pertaining to the Malawi context. We also found that the set of items was slightly skewed towards the “difficult” side of the ability scale. Thus, by removing some of the more difficult items, we were able to group the set of items into two adapted forms—Form A and Form B—that can be used for further studies in the Malawi school context.

References

- Adler, J., & Patahuddin, S. (2012). Recontextualising items that measure mathematical knowledge for teaching into scenario based interviews: an investigation. *Journal of Education*, 56, 17–43.
- An, S., Kulm, G., & Wu, Z. (2004). The pedagogical content knowledge of middle school, mathematics teachers in China and the U.S. *Journal of Mathematics Teacher Education*, 7, 145–172.
- Andrews, P. (2011). The cultural location of teachers’ mathematical knowledge: Another hidden variable in mathematics education research. In T. Rowland & K. Ruthven (Eds.), *Mathematical Knowledge in Teaching* (pp. 99–118). New York, NY: Springer Publications.
- Ball, D. L., & Bass, H. (2003). Toward a practice-based theory of mathematical knowledge for teaching. In B. Davis & E. Simmt (Eds.), *Proceedings of the 2002 annual meeting of the Canadian Mathematics Education Study Group* (pp. 3–14). Edmonton, Alberta, Canada: Canadian Mathematics Education Study Group (Groupe Canadien d’étude en didactique des mathématiques).
- Ball, D. L., & Hill, H. C. (2008). *Mathematical knowledge for teaching (MKT) measures. Mathematics released items 2008*. Ann Arbor, MI: University of Michigan.
- Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, 59(5), 389–407.
- Cole, Y. (2012). Assessing elemental validity: the transfer and use of mathematical

- knowledge for teaching measures in Ghana. *ZDM—The International Journal on Mathematics Education*, 44, 415–426.
- Davis, B., & Simmt, E. (2006). Mathematics-for-teaching: An ongoing investigation of the mathematics that teachers (need to) know. *Educational Studies in Mathematics*, 61(3), 293–319.
- Delaney, S. (2012). A validation study of the use of mathematical knowledge for teaching measures in Ireland. *ZDM—The International Journal on Mathematics Education*, 44, 427–441.
- Delaney, S., Ball, D., Hill, H., Schilling, S., & Zopf, D. (2008). “Mathematical knowledge for teaching”: Adapting U.S. measures for use in Ireland. *Journal of Mathematics Teacher Education*, 11(3), 171–197.
- Edwards, M. C. (2009). An introduction to Item Response Theory using the need for cognition scale. *Social and Personality Psychology Compass*, 3(4), 507–529.
- Fauskanger, J., Jakobsen, A., Mosvold, R., & Bjuland, R. (2012). Analysis of psychometric properties as part of an iterative adaptation process of MKT items for use in other countries. *ZDM—The International Journal on Mathematics Education*, 44, 387–399.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage Publications.
- Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers’ mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, 42(2), 371–406.
- 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.
- Hoover, M., Mosvold, R., & Fauskanger, J. (2014). Common tasks of teaching as a resource for measuring professional content knowledge internationally. *Nordic Studies in Mathematics Education*, 19(3–4), 7–20.
- Kasoka, D., Kazima, M., & Jakobsen, A. (forthcoming 2016). Adaptation of mathematical knowledge for teaching number concepts and operations measures for use in Malawi. To appear in *Proceedings of the 24st Annual meeting of the Southern African Association for Research in Mathematics, Science and Technology, 12th–15th January, 2016*. Pretoria, South Africa: SAARMSTE.
- Kazima, M., & Mussa, C. (2011). Equity and quality issues in mathematics education in Malawi schools. In B. Atweh, M. Graven & P. Valero (Eds.), *Managing Equity and Quality in Mathematics Education* (pp. 163–176). New York, NY: Springer Publications.
- Kwon, M., Thames, M. H., & Pang, J. (2012). To change or not to change: Adapting mathematical knowledge for teaching (MKT) measures for use in Korea. *ZDM—Mathematics Education*, 44, 371–385.
- Malawi Institute of Education (2010). *Initial Primary Teacher Education Programme Handbook*. Domasi, Malawi: Malawi Institute of Education.
- Malawi Institute of Education. (2014). *Improved teaching and learning methods handbook*. Domasi, Malawi: Malawi Institute of Education.

- Malda, M., Van de Vijver, F. J. R., Srinivasan, K., Transler, C., Sukumar, P., & Rao, K. (2008). Adapting a cognitive test for a different culture: An illustration of qualitative procedures. *Psychology Science Quarterly*, 50(4), 451–468.
- Ng, D. (2012). Using the MKT measures to reveal Indonesian teachers' mathematical knowledge: Challenges and potentials. *ZDM—The International Journal on Mathematics Education*, 44, 401–413.
- Rowan, B., Correnti, R., & Miller, R. J. (2002). What large-scale, survey research tells us about teacher effects on student achievement: Insights from the Prospects study of elementary schools. *Teachers College Record*, 104(8), 1525–1567.
- Rowland, T., Huckstep, P., & Thwaites, A. (2005). Elementary teachers' mathematics subject knowledge: The knowledge quartet and the case of Naomi. *Journal of Mathematics Teacher Education*, 8(3), 255–281.
- Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14.
- Wright, S. P., Horn S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education*, 11, 57–67.
- Zimowski, M. F., Muraki, E., Islevy, R. J., & Bock, R. D. (2003). *BILOG-MG 3 for Windows: Multiple-group IRT analysis and test maintenance for binary items [Computer software]*. Lincolnwood, IL: Scientific Software International, Inc.

Acknowledgements

This work was kindly funded by Norwegian Programme for Capacity Building in Higher Education and Research for Development (NORHED) through the project: *Improving quality and capacity of mathematics teacher education in Malawi—a collaboration between University of Malawi, Malawi and University of Stavanger, Norway*.