Assessment of Teacher Knowledge Across Countries: A Review of the State of Research

Sigrid Blömeke and Séan Delaney

Abstract This review presents an overview of research on the assessment of mathematics teachers' knowledge as one of the most important parameters of the quality of mathematics teaching in school. Its focus is on comparative and international studies that allow for analyzing the cultural dimensions of teacher knowledge. First, important conceptual frameworks underlying comparative studies of mathematics teachers' knowledge are summarized. Then, key instruments designed to assess the content knowledge and pedagogical content knowledge of future and practicing mathematics teachers in different countries are described. Core results from comparative and international studies are documented, including what we know about factors influencing the development of teacher knowledge and how the knowledge is related to teacher performance and student achievement. Finally, we discuss the challenges connected to cross-country assessments of teacher knowledge and we point to future research prospects.

Keywords Mathematics teacher education · Comparative study · Mathematics content knowledge (MCK) · Mathematics pedagogical content knowledge (MPCK) · Large-scale assessment

This chapter on the "Assessment of teacher knowledge across countries" brings together two research fields that have developed largely independently from each other: comparative studies on prospective mathematics teachers' content knowledge and pedagogical content knowledge at the end of teacher preparation (Blömeke and Kaiser), and research on the mathematical knowledge of practicing teachers in several countries. The combination of these two research fields provides the reader with

Originally published in ZDM—The International Journal of Mathematics Education (adjusted version for the purpose of this book).

S. Blömeke (✉)
Humboldt-Universität zu Berlin, Berlin, Germany
e-mail: sigrid.bloemeke@staff.hu-berlin.de

S. Delaney
Marino Institute of Education, Dublin, Ireland

S. Blömeke et al. (eds.), *International Perspectives on Teacher Knowledge, Beliefs and Opportunities to Learn*, Advances in Mathematics Education,
DOI 10.1007/978-94-007-6437-8_25,

an overview of what is going on in mathematics education research across countries in relation to teachers' knowledge.

In the first review of this topic for this journal, the state of research was summarized as follows: "Teacher-education research lacks a common theoretical basis, which prevents a convincing development of instruments and makes it difficult to connect studies to each other" (Blömeke et al. 2008a). Since then, research on prospective and practicing mathematics teachers' knowledge has continued to develop. Two research groups have been particularly productive in assessing teacher knowledge with direct measures: one from Michigan State University in the context of "Mathematics Teaching in the 21st Century" (MT21; see, e.g., Schmidt et al. 2011) and the "Teacher Education and Development Study: Learning to Teach Mathematics" (TEDS-M; see, e.g., Blömeke et al. 2012a; Tatto et al. 2012); the second one from the University of Michigan in the context of "Learning Mathematics for Teaching" (LMT; see, e.g., Delaney et al. 2008; Hill et al. 2008). This pioneering work has paved the way for a 2012 ZDM special issue (see, e.g., the papers by Cole 2012; Kwon et al. 2012).

Our chapter summarizes what we have learned in editing this issue, and presents an overview of research on the assessment of mathematics teachers' knowledge going on in other research groups.[1] Our focus is primarily on comparative and international studies because these are innovative areas covering long-lasting research gaps.[2] During the past two decades, the interest in international comparative studies on mathematics teachers has increased (Cochran-Smith and Zeichner 2005; Darling-Hammond 2000). Mathematics teachers will play a central role in the preparation of future generations' K-12 students. An examination of mathematics teachers' knowledge is therefore an important parameter of school quality. It is important to ascertain whether and how teacher training contributes to the development of teacher knowledge.

At the same time, international comparisons allow for analyzing cultural dimensions of teacher knowledge. By developing international studies, many matters are questioned which may remain unquestioned in national studies. The structure and the content of mathematics teacher training depend on a deeper rationale which is a result of factors which may be at least partly cultural. Like the water in the fish's tank, such cultural givens are too often invisible (Blömeke and Paine 2008)—and international comparisons provide the chance to move beyond the familiar, and to see with a kind of "peripheral vision" (Bateson 1994).

The results of comparative studies also provide benchmarks of what level and quality of teacher knowledge can be achieved and point to country-specific strengths and weaknesses. In many countries, the results of such studies on K-12 student achievement have led to fundamental reforms of the school system. The publication of the PISA 2000 results in Germany, for example (Baumert et al. 2001), one of the

[1]We highly appreciate the support of the reviewers who pointed to work relevant for the purpose of our paper. Only due to their efforts were we able to include conceptual frameworks and research from all continents.

[2]We confined our review to articles that were written in the English language.

first international studies the country took part in, and the realization that Germany performed at only a mediocre level—in contrast to the country's self-image—came as a shock. Heated debates and soul-searching among policymakers, researchers, and lay people finally resulted in changes. Similarly, the USA implemented significant reforms in its mathematics school curricula after the so-called "Sputnik shock" and the country's weak performance was confirmed in comparative studies such as SIMS (Pelgrum et al. 1986) and TIMSS (Mullis et al. 1997). Thus, comparative studies of student knowledge provided the chance to understand educational phenomena in a new way. We hope that research on teaching and teacher training across countries will produce similar effects.

Warnings that the proficiency level of mathematics teachers may not be strong enough, given the marginalized role mathematics had been playing in teacher education in many Western countries, had been put forward long ago. Mathematics educators (see, e.g., Schoenfeld 1994; Kilpatrick et al. 2001) and mathematicians (see, e.g., Cuoco 2001; Wu 1999) have repeatedly pointed to the risks of weak training in mathematics: teachers' limited understanding of what mathematics actually is, a fragmented conception with vertical and horizontal disconnects, less than enjoyable teaching routines, and an inability to implement the modern ideas in new school curricula. However, systematic evidence supporting these claims with respect to teachers has been missing for a long time.

Efforts to fill existing research gaps have been made since the late 1990s. Several comparative small-scale studies on mathematics teachers and mathematics teacher training are available (e.g., An et al. 2004; Ma 1999; Burghes 2008). An important step was also the work of a Topic Study Group on mathematics knowledge for teaching at ICME-11 in Mexico (Adler and Ball 2009). About 50 colleagues from a broad range of countries presented their approaches to measuring teachers' or future teachers' mathematical knowledge (see, e.g., Kristjánsdóttir 2008 in several Nordic countries; Naik 2008 in India; West 2008 in Japan). Much of the teacher research, however, neglected the content domain, focused on other subdomains of mathematics teachers' competencies such as beliefs (Bramald et al. 1995; Calderhead 1996), or intended to capture knowledge by self-reports. Studies including direct measures of teacher knowledge and cross-country studies are still needed (Brouwer 2010; Wilson et al. 2001).

Our paper reports about such measures and summarizes their results. It is organized as follows. First, we summarize important conceptual models underlying comparative and international studies of mathematics teachers' knowledge. Second, we develop a model of factors assumed to influence the development of teacher knowledge during teacher training and the relationship between teacher knowledge and student achievement. Third, we describe the study design and key instruments developed to assess the content knowledge and pedagogical content knowledge of future and practicing mathematics teachers. Fourth, core results are documented from comparative and international studies from several countries on the structure and the level of this knowledge and how it is associated with teacher training, teacher performance, and student achievement. Fifth, we discuss the challenges connected to cross-country assessments of teacher knowledge and identify research needed to address these. Each section is subdivided into research on future teachers and practicing teachers.

1 Modeling Mathematics Teacher Knowledge: Conceptual Frameworks

Concepts of teacher knowledge and how it is gained have changed over the past 30 years. A first important model that characterized pre-service teacher education can be labeled as "teacher learning". This included approaches including learning by observation in a kind of apprenticeship (Zeichner 1980), learning by planning, application, and reflection (Schön 1983), and teacher learning as a craft (Brown and McIntyre 1983). The concept's starting point for modeling teachers' competencies was teachers' existing classroom practices.

Similar to this concept was a second one, prominent in the 1990s, in which the cognitive basis of teachers' pedagogical practices started to emerge. The first small-scale comparative studies based on this concept were carried out in the field of mathematics teaching (see especially Pepin 1999; Kaiser 1995).

More recently, teacher education research and research on practicing teachers has focused even more strongly on the knowledge base of teachers' classroom practice. Besides the studies already mentioned in our introduction and the Topic Study Group at ICME-11, this new research paradigm included, for example, studies by Rowland et al. (2005), Chick et al. (2006), and the many chapters from different contexts in Rowland and Ruthven (2010).

Similar to this approach, but more analytical in the sense of defining and distinguishing between different knowledge facets functional for teaching and stressing the importance of mathematics *content* knowledge, is the most recent approach which underpinned not only the PISA study on student achievement but also the two most prominent international teacher studies TEDS-M and LMT. This approach is characterized by a notion of teacher competencies elaborated with respect to the field of mathematics by, for example, Niss (2002) and proficiency in teaching mathematics (see, e.g., Schoenfeld and Kilpatrick 2008).

Elaborating this latter approach, Schoenfeld (2010, p. 187), in his recent book *How we think*, describes in-the-moment decision making as follows: "People's decision making in well practiced, knowledge-intensive domains can be fully characterized as a function of their orientations, resources, and goals." Teaching is such a knowledge-intensive domain, and Schoenfeld points out that mathematics knowledge is the most important resource of mathematics teachers. He regards it as an important enterprise (p. 203) to develop analytical distinctions of knowledge facets and to clarify why particular knowledge facets are accessed in some classroom situations by teachers and others are not. Affective-motivational facets such as orientations and goals or self-regulation are supposed to be decisive in these processes as well. Only together can the full range of teacher competencies underlying classroom performance be understood.

In the following, we present two conceptual frameworks that model the knowledge of mathematics teachers in such a way: the comparative and international studies TEDS-M and LMT.

1.1 Conceptual Framework of TEDS-M as a Comparative Study on Prospective Teachers

In 2008, a comparative study was carried out that examined the knowledge and the beliefs of mathematics teachers at the end of their training: the “Teacher Education and Development Study: Learning to Teach Mathematics” (TEDS-M). Nationally representative samples of primary and lower secondary mathematics teachers in their final year of teacher training from 16 countries in Africa, the Americas, Asia, and Europe were examined, as well as representative samples of teacher educators and training institutions (Tatto et al. 2008). TEDS-M was carried out under the supervision of the International Association for the Evaluation of Educational Achievement (IEA).[3] The study looked at how teachers of mathematics were trained and what kinds of knowledge and beliefs they had at the end of their training with standardized testing. More than 23,000 prospective teachers were surveyed.

The professional knowledge of teachers can be divided into several facets: content knowledge, pedagogical content knowledge, curricular knowledge, and generic pedagogical knowledge (Shulman 1985). In the context of TEDS-M as a study about prospective mathematics teachers, the content knowledge was the knowledge of mathematics. Pedagogical content knowledge referred to knowledge about the teaching and learning of mathematics. In the TEDS-M framework, it included curricular knowledge, too. Pedagogical knowledge, finally, was the knowledge typically acquired in a teacher training program that is not subject-matter related (Blömeke and Paine 2008).

Mathematics content knowledge (MCK), in this framework, includes not only basic factual knowledge of mathematics but also the conceptual knowledge of structuring and organizing principles of mathematics as a discipline (Shulman 1987): why a specific approach is important and where it is placed in the universe of approaches to mathematics. *Mathematics pedagogical content knowledge* (MPCK) includes the subject-related knowledge for teaching. Shulman (1987, p. 9) characterizes it as an “amalgam of content and pedagogy that is uniquely the province of teachers, their own special form of professional understanding”. A mathematics teacher has to know about typical preconditions of students and how to represent a topic in the best possible way. Curricular knowledge is part of it and includes teaching materials and curricula.

[3]TEDS-M was funded by the IEA, the National Science Foundation (REC 0514431), and the participating countries. In Germany, the German Research Foundation funded TEDS-M (DFG, BL 548/3-1). The instruments are copyrighted by the International Study Center at Michigan State University (ISC). The views expressed in this paper are those of the authors and do not necessarily reflect the views of the IEA, the ISC, the participating countries, or the funding agencies.

Table 1 Core situations which mathematics teachers are expected to manage (Tatto et al. 2008)

Mathematical curricular knowledge	Establishing appropriate learning goals
	Knowing about different assessment formats
	Selecting possible pathways and seeing connections within the curriculum
	Identifying the key ideas in learning programs
	Knowledge of the mathematics curriculum
Knowledge of planning for mathematics teaching and learning [pre-active]	Planning or selecting appropriate activities
	Choosing assessment formats
	Predicting typical student responses, including misconceptions
	Planning appropriate methods for representing mathematical ideas
	Linking didactical methods and instructional designs
	Identifying different approaches for solving mathematical problems
	Planning mathematics lessons
Enacting mathematics for teaching and learning [interactive]	Analyzing or evaluating students' mathematical solutions or arguments
	Analyzing the content of students' questions
	Diagnosing typical student responses, including misconceptions
	Explaining or representing mathematical concepts or procedures
	Generating fruitful questions
	Responding to unexpected mathematical issues
	Providing appropriate feedback

An important implication of the TEDS-M framework that modeled teacher knowledge as a facet of teacher "competencies" (Weinert 2001) was its link to classroom situations. Since competencies were assumed to represent a latent trait that underlies performance, teacher knowledge was regarded as situated and applied by nature (Blumer 1969). The classroom situations a teacher has to deal with have to be set by constitutive features of the teaching profession. To determine which features were to be regarded as constitutive, TEDS-M referred to existing standards for the national teacher training programs in its participating countries (e.g., KMK 2004; NCTM 2000). Table 1 documents the problems which mathematics teachers were expected to solve in TEDS-M according to the study's conceptual framework (see Tatto et al. 2008), based on these standards.

In this applied perspective, it is important to distinguish between different types of MPCK (Shulman 1986): knowledge in teaching, that is *propositional knowledge* of, for example, student errors or misconceptions without being related to a specific classroom context; *case-based knowledge* that includes prototypes, borderline cases, and analogies based on individual experiences; and *strategic knowledge* or "practical wisdom" for situations when a teacher is overwhelmed by the multidimensionality and speed of what is going on in the classroom. As far as is possible

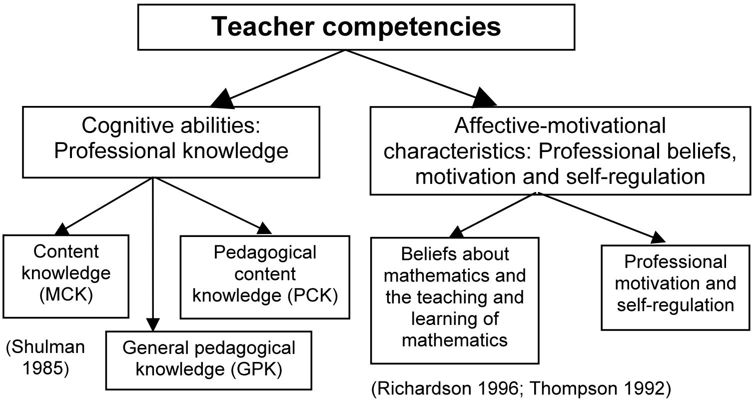


Fig. 1 Conceptual framework of teacher competencies

in a paper-and-pencil test, TEDS-M tried to cover the first two types of knowledge, the propositional and the case-based knowledge of mathematics teachers.

TEDS-M looked also at the professional beliefs held by the future mathematics teachers. Beliefs were defined by Richardson (1996, p. 103) as "psychologically held understandings, premises, or propositions about the world that are felt to be true". As Schoenfeld (2010) pointed out, teacher beliefs are crucial for the perception of classroom situations and for decisions on how to act (Leder et al. 2002; Leinhardt and Greeno 1986). Therefore, they connect knowledge and action. In this sense, they are also an indicator of the type of instruction that mathematics teachers may use in their future teaching (Brown and Rose 1995). If beliefs are operationalized specifically to both the content being taught and the challenges a specific classroom situation presents, empirical evidence exists for a link between teacher beliefs and student achievement (Bromme 1994). In TEDS-M, several beliefs facets were distinguished, in particular epistemological beliefs about the nature of mathematics and beliefs about the teaching and learning of mathematics (Thompson 1992).

By distinguishing between knowledge and beliefs and thus to include cognitive and affective-motivational teacher dispositions and by stressing the situative and applied nature of teacher knowledge, the TEDS-M framework can be connected to a measurement tradition prominent in educational psychology: the measurement of competencies. Competencies as defined in general by Weinert (2001) and specifically with regard to teaching by Bromme (1997) mean having the cognitive ability to develop effective solutions for job-related problems and, in addition, having the motivational, volitional and social willingness to successfully and responsibly apply these solutions in various situations (see Fig. 1). The purpose of such a multidimensional approach is to come as closely as possible to real behavior in the classroom that is supposed to be guided by both types of dispositions. Measuring them with separate instruments allows for examining the relationship and clustering teachers according to their beliefs-knowledge profile (Blömeke et al. 2012a).

Besides beliefs as one facet of the affective-motivational dimension, a teacher's professional motivation and self-regulation are important. A teacher who regulates her behavior is able to define her professional objectives, to decide on appropriate strategies in order to achieve her objectives, and to apply them in various situations. Furthermore, she monitors and evaluates her behavior systematically guided by metacognition (Butler and Winne 1995; Perry et al. 2006; Boekaerts and Corno 2005). Teaching strengths and weaknesses are identified and—if necessary—either behavior is adjusted or professional development activities are taken. Self-regulation capacities are a necessary precondition in order to be successful as a teacher in the long run.

Such a multidimensional notion of teacher competencies that underlie classroom performance led not only the item development in TEDS-M but also comparative studies such as MT21 and the follow-up studies to TEDS-M (TEDS-FU) carried out in Germany, Taiwan, and the USA. TEDS-FU currently looks into the transition of the samples tested in TEDS-M from teacher training into the job (see the forthcoming special issue of the *International Journal of Science and Mathematics Education*, *IJMA*, edited by the three national TEDS-M coordinators in these countries, Blömeke, Schmidt, and Hsieh).

However, such an analytical approach has also been criticized. It has its benefits with respect to the purpose of reliable measurements, but from a didactical perspective that focuses on the practical side of a teacher's work it is argued that the separate dimensions have to come together in the classroom (Huillet 2009; Bednarz and Proulx 2009). This is certainly a valid observation. We will come back to it at the end of our paper when we discuss issues of validity in our efforts to measure teacher knowledge.

1.2 Conceptual Framework of Studies on Practicing Teachers

The studies on practicing teachers that are considered here differ from the TEDS-M studies in at least two ways. First, the studies focus mainly on elementary teachers. Although measures of secondary teachers' mathematical knowledge which can be used at scale have been developed by the German COACTIV (Cognitive Activation in the Classroom) research group (e.g., Krauss et al. 2008a), none of the studies reported here have adapted these measures for use in other countries. Instead, the measures used were developed for use with elementary teachers by the US LMT (Learning Mathematics for Teaching) research group (e.g., Hill and Ball 2004). Although the LMT group has subsequently developed measures for studying teachers of middle school students, the team's original focus was on knowledge held by teachers of elementary students.

Second, unlike the conceptual framework which was developed for the purposes of comparing prospective teachers participating in the TEDS-M study, the conceptual framework for mathematical knowledge for teaching (MKT) was developed by researchers at the University of Michigan (Ball and Bass 2003, p. 399) in order to better understand teacher knowledge in the United States. The conceptual framework of MKT was inspired by Shulman's (1986) idea of pedagogical content knowledge and seeks to categorize the domains of knowledge needed to do the work of teaching mathematics (Ball et al. 2008). It consists of two broad categories—subject matter knowledge and pedagogical content knowledge—and does not include affective-motivational characteristics. Each broad category has been further subdivided into the domains of common content knowledge, specialized content knowledge, and horizon content knowledge on one hand; and knowledge of content and students, knowledge of content and teaching, and knowledge of content and curriculum on the other. Each domain refers to a hypothesized type of mathematical knowledge that is needed by teachers.

Common content knowledge (CCK) refers to mathematical knowledge "used in settings other than teaching" (Ball et al. 2008, p. 399) and an example would be recognizing and naming a two-dimensional shape such as a rectangle or a pentagon. Specialized content knowledge (SCK) is mathematical knowledge and skill that is "not typically needed for purposes other than teaching" (Ball et al. 2008, p. 400), such as knowing a range of definitions of shapes that are both comprehensible to students of different age levels, and mathematically accurate and complete. Knowledge of content and students (KCS) "combines knowing about students and knowing about mathematics" (Ball et al. 2008, p. 401) and would involve knowing for example that a square remains a square even if it is rotated 45 degrees. Knowledge of content and teaching (KCT) refers to knowledge of mathematics combined with knowledge of teaching and would include knowing how to select a poster to support the teaching of shapes by using non-examples and non-stereotypical examples. A provisional domain is horizon content knowledge (HCK), which is an "awareness of how mathematical topics are related over the span of mathematics included in the curriculum" (Ball et al. 2008, p. 403). HCK includes knowledge of the wider discipline of mathematics insofar as its content and practices can inform the work of teaching and, as a domain, requires further specification (Jakobsen et al. 2012). Knowledge of content and curriculum (KCC) is another provisional category. Empirical evidence for these domains has been sought but findings to date have been inconsistent (Schilling et al. 2007). Measures based on the framework have been developed for all domains except HCK and KCC and it is these measures which have appealed to researchers outside the United States.

An important distinction to notice between the TEDS-M study and studies using measures based on MKT is that the TEDS-M study was conducted in conjunction with the IEA, whereas the theory of MKT and the measures based on the theory were developed by the US-based Learning Mathematics for Teaching (LMT) project. Although both frameworks are related to classroom situations, the teaching which informed the development of MKT was specifically US teaching (Ball et al. 2008). Consequently, when measures based on the theory are used outside the United States, researchers need to evaluate the suitability of using US measures to study the mathematical knowledge held by teachers in other countries. This is because the theory is based on the practice of teaching, and if the practice of teaching is a cultural activity as some would argue (e.g. Stigler and Hiebert 1999) then the theory of MKT may also be culture-specific. Nevertheless, the availability of high quality measures of MKT that could be used at scale (Hill et al. 2004) has made it worthwhile to evaluate their suitability for use in several countries, despite the challenges involved in doing so.

One challenge is that a coordinated study of practicing teachers' mathematical knowledge across countries has not yet been funded. This situation contrasts with the substantial funding of several large-scale comparative studies—TIMSS and PISA—of students' mathematical knowledge. Studies which have used the MKT framework to study teacher knowledge, including those published here, tend to be conducted locally, are small-scale in nature, and are situated in specific countries. However, despite being small in scale, the lessons learned in such studies can inform comparative studies of teacher knowledge, should they take place in the future. The range of settings for the studies—involving countries in Europe (Ireland and Norway), Africa (Ghana), and Asia (Indonesia and Korea)—provides a robust and diverse test for the US measures. Such diversity is likely to ensure that the studies are well placed to evaluate as well as to contribute to the conceptual framework of MKT.

In addition to the studies published in the 2012 ZDM special issue (see, e.g., the papers by Cole 2012; Kwon et al. 2012), the MKT approach and framework (but not the measures) have been used as the basis of research into teacher knowledge by scholars in settings as diverse as New Zealand (Burgess 2009) and South Africa (Adler and Davis 2006; Kazima and Adler 2006). In addition to studies of practicing teachers, the framework has been used in studies of mathematical knowledge among Australian and Canadian pre-service teachers (Butterfield and Chinnappan 2010, 2011). Similarly, Forrester and Chinnappan (2011) studied 224 pre-service teachers in a year 1 subject in Bachelor of Education Primary drawing on the work of Ball and Bass (2003).

Alternative frameworks for studying teachers' mathematical knowledge have also been proposed. Chick and her colleagues developed a framework that includes categories that are "clearly PCK", "content knowledge in a pedagogical context", and "pedagogical knowledge in a content context" (e.g., Chick et al. 2006). Also in Australia, Chinnappan and Lawson (2005) developed a framework for studying knowledge of geometry for teaching and applied it to studying the knowledge exhibited by two experienced teachers in taped interviews. One study that did not look at teachers but at the history of mathematics, research studies, and textbooks is the work in Brazil of Ribeiro (2008) who elaborates on how understanding six specified meanings of equations can contribute to teachers' knowledge for teaching.

At the secondary level, with the exception of the COACTIV work, frameworks for studying mathematical knowledge for teaching are a more recent development. However, researchers at Penn State University have analyzed the mathematics needed by teachers of secondary students from three perspectives: a curriculum-based mathematical thinking approach, a situations approach, and a mathematical process approach (Heid 2008). A series of studies have been designed by Penn State faculty and graduate students around these perspectives to characterize prospective teachers' mathematical knowledge, to characterize their knowledge and use of mathematical processes, and (in conjunction with faculty and graduate students from the

University of Georgia) to develop a practice-based framework of MKT at the secondary level (Heid 2008).

Like most studies of teachers' mathematical knowledge, the examples we have found tend to be conducted locally and are small in scale. We found more studies of primary teachers than of secondary teachers. Many studies looked at pre-service teachers, probably because as subjects they are much more accessible to university-based researchers. Because primary teachers tend not to have subject-specific preparation, primary teachers have long been perceived by many to be lacking in MKT. However, studies of prospective and practicing secondary teachers are appearing more frequently than before as researchers find that specialized subject preparation does not necessarily equip teachers with the knowledge they need for teaching.

2 Modeling the Link from Mathematics Teacher Education to Student Achievement

The topic of this book and our review, the knowledge of prospective and practicing mathematics teachers, is one[4] crucial link between mathematics teacher education and student achievement in mathematics. In the first perspective, teacher knowledge represents a core criterion for effective teacher training and is thus a dependent variable. In the second perspective, teacher knowledge represents an important predictor of student achievement and is thus an independent variable. Only both perspectives together provide an appropriate view on the relationship of mathematics teacher training and what is accomplished in schools—multiply mediated by complex context factors.

Before we start to model this link based on the frameworks used in TEDS-M and LMT, we have to point out that the enterprise represents a huge—and maybe at present too huge—step. We have currently only data from studies examining parts of the connection between mathematics teacher education, mathematics teacher knowledge, and student achievement. Thus, we cannot claim that the overall model stands on strong grounds. Further, the studies are relational only and not longitudinal, thus causal claims cannot be made, either. This section is therefore meant to build a heuristic rather than to give a definite answer.

2.1 Mathematics Teacher Education and Teachers' Professional Knowledge

In order to examine which factors may influence the development of teacher knowledge during teacher training, potentially influential factors were divided into three categories in TEDS-M:

[4]From here on, we focus on the state of research on teacher knowledge that is related to mathematics. This topic is already highly complex in a comparative context and covering pre-service and in-service teachers. General pedagogical knowledge, beliefs, motivation, and self-regulation have thus to be neglected although they are also highly important for teaching (for results on the future teacher's belief and their relationship to knowledge see, e.g., Blömeke et al. 2012b).

(1) the individual characteristics of future teachers
(2) the institutional characteristics of teacher training, and
(3) the national country context.

In studies of school effectiveness, K-12 students' background is almost always a powerful predictor of achievement. Specifically with respect to mathematics, gender (Hyde et al. 2008), socio-economic status (Mueller and Parcel 1981), and language background (Walter and Taskinen 2008) as well as prior knowledge (Simmons 1995) and motivation (Singh et al. 2002) play an important role (Scheerens and Bosker 1997). Such a network of individual predictors may apply to the knowledge acquisition during mathematics teacher training as well.

With respect to institutional characteristics, TEDS-M followed the tradition of the IEA in connecting educational opportunity and educational achievement. As was done in the "Third International Mathematics and Science Study" (TIMSS), opportunities to learn (OTL) were framed as content coverage, specifically as "the content of what is being taught, the relative importance given to various aspects of mathematics and the student achievement relative to these priorities and content" (Travers and Westbury 1989). OTL were in this sense defined as future mathematics teachers' encountering occasions to learn about particular topics during teacher training, including the characteristics of their educators such as background, teaching objectives, and teaching methods used. Since subject matter specificity is the defining element of an educational opportunity (Schmidt et al. 1997), the particular topics reflected the areas of mathematics and mathematics pedagogy.

OTL in teacher training can be regarded as having been intentionally developed by educational policy makers and teacher training institutions (Stark and Lattuca 1997; Schmidt et al. 2008). They give characteristic shape and direction to instruction. Every choice provides some OTL at the expense of others. National program choices in this sense reflect particular visions of what mathematics teachers are supposed to know and be able to do in class and how teacher training should be organized in order to provide the knowledge and skills necessary for successful accomplishment of their professional tasks.

Thus, teacher knowledge at the end of teacher training is assumed to depend on individual background characteristics of the prospective mathematics teachers, their differential learning experiences during teacher training, opportunities to learn provided by their training institutions, and the social, schooling, and policy context they are living in (see Fig. 2). However, as a study on the effectiveness of teacher education, TEDS-M did not look into the relationship between teachers' professional knowledge and student achievement.

2.2 *Mathematics Teachers' Professional Knowledge and Student Achievement*

While the characteristics of prospective mathematics teachers depend on their entering characteristics, the learning experiences received from interactions with teacher

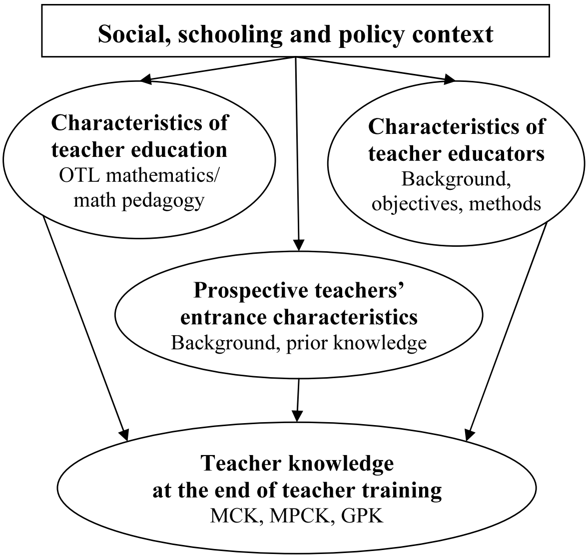


Fig. 2 Hypothetical model of characteristics of teacher training influencing teacher knowledge mediated by the prospective teachers' entrance characteristics (based on Tatto et al. 2008)

educators and their teacher education program, and from their social, schooling, and policy context, additional factors such as experience (e.g. Hill 2007) and professional development (Bell et al. 2010) may influence the knowledge held by practicing teachers. Our interest, however, in teachers' knowledge is not ultimately an end in itself but as a means of improving student achievement.

For many years researchers have attempted to study the link between teacher mathematical knowledge and student achievement in what are often referred to as educational production function studies (Monk 1989). Begle conducted an early educational production function study in 1972 (Begle 1972; Eisenberg 1977). He subsequently used his own research and studies by others to conclude that, beyond a certain level, mathematical knowledge matters little for student achievement (Begle 1979). In the studies referred to by Begle, mathematical knowledge was measured by tests of general mathematical knowledge administered to teachers (e.g. Begle 1972). Begle's studies and the studies he reviewed provided evidence that proxy measures of teacher knowledge (e.g., math courses studied) and performance on generic mathematics test items are not good predictors of student learning, suggesting that more sophisticated means of studying teachers' mathematical knowledge were needed.

More recent studies that attempt to link teacher knowledge and student achievement have used "pencil-and-paper" tests. Their goal is to test teachers' performance on mathematics problems that occur in teaching rather than on general mathematical knowledge tasks. They focus on specific mathematical domains, and have been developed by teams with expertise in mathematics, mathematics education, and psychometrics (Hill et al. 2007b). An example of such a study is one by Hill et al. (2005)

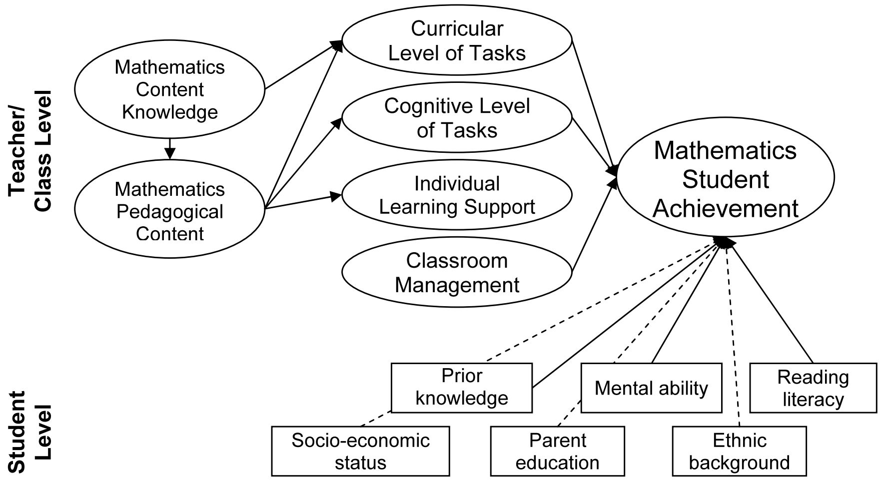


Fig. 3 Hypothetical model of the effects of teachers' mathematics content knowledge and mathematics pedagogical content knowledge on student achievement in mathematics mediated by instructional quality (based on Baumert et al. 2010)

where it was found that every standard deviation of difference in teacher knowledge was worth the equivalent of two to three weeks additional instruction time in predicting the gains made by first and third grade students on standardized math test scores. In third grade the effect size of teacher knowledge "rivaled that of SES and students' ethnicity and gender, while in the first grade models the effect size was not far off" (p. 396). Although such findings in a single study need to be treated with caution, they confirm the importance of pursuing teacher knowledge as a key variable in student achievement.

A longitudinal study conducted by Baumert and colleagues (2010) in Germany studied the effects of one domain of mathematical knowledge, pedagogical content knowledge (PCK), on student achievement at the end of grade ten (see Fig. 3). The study found a substantial effect size of PCK on student achievement, an effect which is distinguishable from the effect of general mathematical knowledge and other factors such as track membership (i.e. membership in one of the different school types existing parallel to each other in Germany). Although this finding too is promising, it is limited to studying teacher knowledge in a single country.

Comparative studies of teacher knowledge have tended to be on a smaller scale and consequently have not established "direct and statistical relationships between teaching-related factors and student mathematics performance" (Wang and Lin 2005). Although studies such as Ma's (1999) demonstrated that some Chinese teachers have deeper mathematical knowledge than some US teachers, another study found that whether US or Chinese students perform better on mathematics problems varies by problem type (Cai 2000). More importantly for our interest, factors other than teacher knowledge, such as classroom instruction or placement of topics on the school curriculum (Cai 2000; Wang and Lin 2005), may account for differences in student performance between the two countries. Studying the relationship between teacher knowledge and student achievement across several countries could yield important insights in mathematics education. Consequently, insights from country-specific studies of the teacher knowledge/student achievement relationship and from studies of teacher knowledge across countries will inform how challenges inherent in such research can be addressed.

3 Study Design and Instruments to Assess Mathematics Teachers' Knowledge

Most comparative studies have to deal with the same methodological challenges. In order to reduce the complexity, this section focuses again on selected studies as examples. The largest study on prospective mathematics teachers was TEDS-M. Its study design and instruments are therefore described and evaluated. An instrument for the assessment of practicing teachers' mathematical knowledge applied in many countries is the LMT test, which is also documented.

3.1 Design of TEDS-M as a Study on Prospective Teachers

3.1.1 Sampling

The target population of TEDS-M consisted of students in their final year of teacher training who were on track to receive a license to teach mathematics in primary or lower secondary schools (Tatto et al. 2008). Since the definition of "primary" and "lower secondary" varies across countries, an agreement on which programs to include in which of the two studies was difficult to accomplish. The complexity was increased by the substantial number of programs covering both school levels. Finally, a teacher training program was included if it prepared primary teachers for one of the grades 1 through 4 as the common denominator of level 1 education in the "International Standard Classification of Education" (primary or basic education, cycle 1; UNESCO 1997), or if it prepared lower secondary teachers for grade 8 as the common denominator of level 2 education (lower secondary or basic education, cycle 2). Programs covering both levels were included in both studies.

In a two-stage process, random samples were drawn from the target populations in each participating country. The samples were stratified according to important teacher training features such as "route" (consecutive vs. concurrent programs), "type" of program (grade span the license includes, e.g., grades 5 through 9 vs. 5 through 12), "focus" of opportunities to learn (with or without extensive mathematics), and "region" (e.g., federal state) in order to reflect accurately the distribution of prospective primary and lower secondary teachers' characteristics at the end of training.

In 2008, approximately 14,000 prospective primary and 9,000 lower secondary mathematics teachers from altogether 16 countries (15 countries in each of the two

Table 2 Countries participating in TEDS-M

Botswana	Chile	Germany	Georgia
Malaysia	Norway	Oman (Lower secondary study only)	Philippines
Poland	Russia	Switzerland	Singapore
Spain (Primary study only)	Taiwan	Thailand	USA

studies; see Table 2) were tested on their knowledge of mathematics and mathematics pedagogy by a standardized paper-and-pencil assessment. Rigorous quality requirements, as set out in TIMSS or PIRLS by the IEA, included controlling the translation processes, monitoring test situations, and meeting the required participation rates. The aim of these standards was to ensure that bias resulting from non-response was kept within acceptable limits. However, in contrast to studies on student achievement where ministries of education often can pledge school and teachers to participate, TEDS-M had to deal with universities which enjoy a large degree of autonomy. Furthermore, future teachers are adults who can decide about participation by themselves. Correspondingly, it was much harder to meet IEA's requirements than it is on the school level.[5]

3.1.2 Instruments

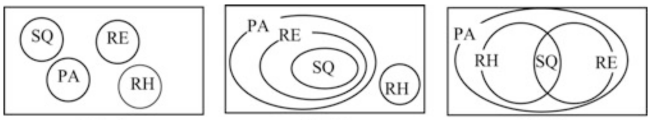
TEDS-M sought to measure prospective teachers' MCK and MPCK at the end of their training. For this purpose, a 60-minute paper-and-pencil assessment was completed during a standardized and monitored test session. The MCK test covered number, algebra, and geometry with approximately equal weight and, to a lesser extent, data (Tatto et al. 2008). In addition, three cognitive dimensions were covered by the items: knowing, applying, and reasoning. Sample items are given in Figs. 4 and 5 (the full sets of released items are available from tedsm@msu.edu).

The MPCK test covered two facets: knowledge of curricula and planning, which is necessary before a teacher enters the classroom, and interactive knowledge about how to enact mathematics for teaching and learning. In line with the MCK test, four content areas were distinguished. An example is given in Fig. 6.

The item development was mainly informed by the MT21 study (Schmidt et al. 2011), as well as the two Michigan studies entitled "Knowing Mathematics for Teaching Algebra" (KAT; Ferrini-Mundy et al. 2005) and "Learning Mathematics for Teaching" (LMT; Hill et al. 2008). Three item formats were used: multiple choice, complex multiple choice, and open constructed response.

[5]The participation rates in four countries on the primary level (Chile, Norway, Poland, and the USA) and five countries on the lower secondary level (Chile, Georgia, Norway, Poland, and the USA) did not fully meet the required benchmarks. Their results are therefore reported in an annotated way. In Poland, Switzerland, and the USA the coverage of the target population was reduced and in Norway the sample composition did not fully meet the TEDS-M definition of the target population. In the USA, a substantial proportion of missing values was observed. The results of these countries are reported in an annotated way as well. Canada had to be excluded from the study because the country missed the benchmarks to a serious extent.

Three students have drawn the following Venn diagrams showing the relationships between four quadrilaterals: rectangles (RE), parallelograms (PA), rhombuses (RH) and squares (SQ).



[Tian] [Rini] [Mia]

Which student's diagram is correct? *Check one box.*

A. [Tian] $\circ_1$

B. [Rini] $\circ_2$

C. [Mia] $\circ_3$

Fig. 4 Sample item from the TEDS-M primary test of MCK

Fig. 5 Sample item from the TEDS-M lower secondary test of MCK

Prove the following statement:
If the graphs of linear functions

$$f(x) = ax + b$$

and

$$g(x) = cx + d$$

intersect at a point *P* on the *x*-axis, the graph of their sum function

$(f + g)(x)$ must also go through *P*.

Fig. 6 Sample item from the TEDS-M primary test of MPCK

When teaching children about length measurement for the first time, Mrs. [Ho] prefers to begin by having the children measure the width of their book using paper clips, and then again using pencils.

Give **TWO** reasons she could have for preferring to do this rather than simply teaching the children how to use a ruler.

3.1.3 On the Nature of the TEDS-M Knowledge Tests

If one is to evaluate the nature of the TEDS-M tests as done by Döhrmann, Kaiser, and Blömeke (this book), one can summarize that the MCK and MPCK of prospective teachers was successfully conceptualized and efficiently surveyed through the TEDS-M assessments. The authors confirm the overall reliability and validity of the tests from an international point of view. However, they also point to limits of the assessment. Due to cultural differences between the participating countries, the items did not cover the entire range of the knowledge teachers should acquire during teacher training. Some teacher tasks are relevant in certain countries but not in others. Thus, the corresponding knowledge was not assessed.

In addition, Döhrmann, Kaiser, and Blömeke (this book) characterize the orientation of the conceptual framework and the item pool as slightly biased towards a pragmatic conception of teaching and learning, predominantly in place in English-speaking countries. Facets common in continental Europe were taken into account to some extent only. In this European tradition, the mathematics and the mathematics pedagogy tests would have had to include more argumentation and proof items, for example. Also, fundamental ideas of central mathematical concepts such as number or percentage and different ways to introduce them in class would have had to have a higher priority.

Hsieh, Lin, and Wang (this book) argue along the same line. Confirming the overall validity of the TEDS-M tests, they point out that the tests have more items testing MCK than MPCK, which may demonstrate an unbalanced focus of interest. Based on conceptual considerations or empirical evidence, many colleagues would make a plea for MPCK as the crucial facet of teacher knowledge (see, e.g., Baumert et al. 2010). Furthermore, Hsieh, Lin, and Wang (this book) would prefer an assessment more closely in line with Niss (2002), who emphasized the fundamental characteristics of mathematical thinking across the content domains of mathematics rather than assessment of their knowledge in these specific domains (see also Hsieh 2010; Hsieh et al. 2010). This perspective is also supported by An and Wu (2011).

An open question is to what extent the TEDS-M tests were measurement invariant across countries. The number of countries was only 15, with even smaller numbers of country groups from similar educational traditions or with substantial proportions of teachers using different languages within the countries. Thus, a potential cultural bias or a potential language bias could only be examined to a limited extent. First evidence suggests that the MCK and MPCK assessments may not have been *completely* equivalent in all TEDS-M countries (Blömeke et al. 2011a). Although rigorous quality control had taken place, language and cultural differences between and within countries seem to be related to how well these traits were measured in the TEDS-M countries. The language problem seemed to be larger with respect to MCK than to MPCK. Blömeke et al. (2011a) attributed this result to a long history of schooling in a different language from that used at home in the case of MCK. Its acquisition had probably already suffered from language disadvantages before the prospective teachers entered university.

3.1.4 Challenges of Scaling the TEDS-M Data

Scaled scores in TEDS-M were created separately for MCK and MPCK in one-dimensional models using item response theory. Such models stress the conceptual difference between the two knowledge facets (see Tatto et al. 2012). An alternative way of scaling the data was used by Blömeke et al. (2011a). They applied two-dimensional models that can take the conceptual overlap between MCK and MPCK into account (Reckase 2009). Whereas traits such as reading literacy or mathematics literacy, typically found in PIRLS or TIMSS, are relatively well-defined, such conceptual clarity does not exist with respect to teacher knowledge. Researchers are still struggling to separate its facets (Graeber and Tirosh 2008). Studies by Schilling et al. (2007) or Krauss et al. (2008b) demonstrated that MCK and MPCK were highly correlated.

Therefore, Blömeke et al. (2011a) used a multidimensional IRT approach in which MCK and MPCK were represented with a general and a nested factor ("within-item multidimensionality", Adams et al. 1997). The model represented Shulman's idea that the nested factor MPCK was a mixture of different abilities and that mathematics pedagogy items measured this mix. According to this idea, solving the mathematics pedagogy items required MCK as a general ability but also specific MPCK. The results supported, in fact, the contention that the nature of teacher knowledge is multidimensional (Blömeke et al. 2011a). Only in this model was the importance of opportunities to learn mathematics pedagogy during teacher education revealed, too. The more a country had focused on mathematics pedagogy in relation to mathematics during teacher training, the more likely it would be to have a high MPCK mean.

3.1.5 Challenges of Reporting the TEDS-M Results

The TEDS-M results gave a sound picture of the professional knowledge of prospective mathematics teachers who in 2008 were in their final year of teacher training. However, there was an important challenge in how to report the results. The data had to be analysed on two levels of aggregation because it was important to distinguish between an evaluation of the national teacher training systems and an evaluation of specific teacher training programs within countries. Both approaches have their benefits and their limitations:

(1) Due to the traditional policy orientation of IEA's large-scale assessments, TEDS-M focused on the national level. This approach stressed the overall educational effectiveness of a nation, regardless of the structure of its education system. In this perspective, with regard to international competitiveness, it considered what a nation accomplishes as a whole—and differences in the structure of teacher education systems between countries represent a function of differences in their educational policy.

(2) Additional information was gained by looking into program types. Thus, it was possible to learn about pathways to success within countries, that is, without confounding variables like cultural or societal features. Note that the relatively small sample sizes in the case of teachers (compared with students) became even smaller when types of programs were examined and that the precision of estimates was probably lower because the sampling target was mainly on the national level. This approach has therefore to be used with caution.

Ms. Chambreaux's students are working on the following problem:

Is 371 a prime number?

As she walks around the room looking at their papers, she sees many different ways to solve this problem. Which solution method is correct? (Mark ONE answer.)

a) Check to see whether 371 is divisible by 2, 3, 4, 5, 6, 7, 8, or 9.
b) Break 371 into 3 and 71; they are both prime, so 371 must also be prime.
c) Check to see whether 371 is divisible by any prime number less than 20.
d) Break 371 into 37 and 1; they are both prime, so 371 must also be prime.

Fig. 7 Sample CCK item from the LMT database of MKT multiple choice items. Released items are available from http://sitemaker.umich.edu/lmt/files/LMT_sample_items.pdf

3.2 Use of MKT Measures to Study Practicing Teachers

Because the studies that used the MKT measures to study teachers' knowledge are country-specific, they differ from each other in terms of how samples of respondents were selected, in terms of the specific measures used, and in terms of how data were analyzed. This is because studies conducted to date have not been coordinated centrally and the purpose and the resources available differed from one setting to another. What they have in common is that they all accessed the database of multiple choice measures of MKT that were developed at the University of Michigan. The measures available related primarily to three domains of MKT: CCK, SCK, and KCS.

Figure 7 is an example of an MKT item which is considered to tap into teachers' common content knowledge (CCK). A teacher responding to the item needs to consider various strategies that were used to check if a number is prime, where only one strategy is valid. Answer (a) is incorrect because it is possible that 371 could be divisible by two numbers higher than 9 and not by a number less than 9. The number 19 is the square root of 371 making 19 a factor (although in this case 7, one of the numbers to be checked, is also a factor). Response option (b) is incorrect because it is invalid to test for primes by breaking down the number additively into hundreds and tens and units. In this specific case, for example, the 3 represents 300 which is not a prime number. Response (c) is correct because 20 is greater than the square root of 371 and if 371 has a prime factor greater than 20, it must have one that is less than 20 as well. Response (d) is incorrect for reasons similar to response (b). This item taps into general mathematical knowledge of prime numbers and although it is information a teacher would use, it is not specific to the work of teaching and it requires no knowledge of students or teaching.

Figure 8 is an example of an MKT measure that taps into specialized knowledge of mathematics. A respondent is expected to analyze four different potential repre-

At a professional development workshop, teachers were learning about different ways to represent multiplication of fractions problems. The leader also helped them to become aware of examples that do <u>not</u> represent multiplication of fractions appropriately.

Which model below <u>cannot</u> be used to show that $1\frac{1}{2} \text{ x } \frac{2}{3} = 1$? (Mark ONE answer.)

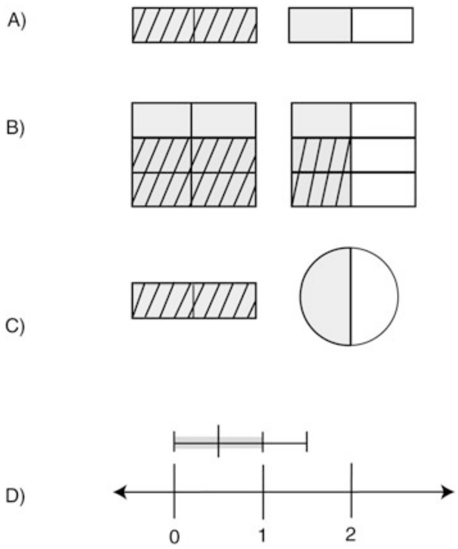
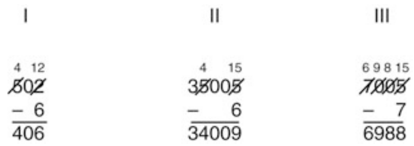


Fig. 8 Sample SCK item from the LMT database of MKT multiple choice items. Released items are available from http://sitemaker.umich.edu/lmt/files/LMT_sample_items.pdf

sentations of a calculation involving multiplication of fractions. In response (a) two unit area models are shown, with 1½ shaded in grey and 2/3 of the 1½ indicated with oblique lines. Response (b) is similar but the units are partitioned into sixths rather than halves. The third model is considered to be unsuitable because one unit is a rectangle and the second is a circle. Response (d) uses a number line model to show a line measuring 1½ partitioned in three equal parts, each measuring 1/2. Two of these parts (or 2/3 of the line measuring 1½ units) are shaded indicating the product of 1½ and 2/3. The knowledge required here is purely mathematical but it is mathematical knowledge that is specific to teaching.

Figure 9 is a sample KCS item and draws on the teacher's knowledge of mathematics and students combined. In three samples of work, students have subtracted incorrectly. However, the errors made differ. In sample I, the student exchanged one hundred for ten units and in sample II the student exchanged one thousand for ten units. A teacher who has knowledge of students and mathematics will recognize that both problems relate to a misunderstanding of renaming a number using principles of place value. In sample III the student exchanged one thousand for nine hundreds, eight tens and ten units, which is also incorrect but the error made is more sophis-

Mrs. Jackson is planning mini-lessons for students around particular difficulties that they are having with subtracting from large whole numbers. To target her instruction more effectively, she wants to work with groups of students who are making the same kind of error, so she looks at a recent quiz to see what they tend to do. She sees the following three student mistakes:



Which have the same kind of error? (Mark ONE answer.)

a) I and II

b) I and III

c) II and III

d) I, II, and III

Fig. 9 Sample KCS item from the LMT database of MKT multiple choice items. Released items are available from http://sitemaker.umich.edu/lmt/files/LMT_sample_items.pdf

ticated than a direct swap of hundreds or thousands for ten units as is the case in I or II, so the correct answer is (a).

A substantial database of MKT items has been developed and many of these have appeared on various survey forms used in studies by the research group. The LMT research team has set a standard for using such items and for analyzing responses. The purposes for which measures may be used and how results are presented are restricted and users of the measures are required to participate in a training workshop before using the measures. They may not be used for high stakes purposes such as decisions related to appointments or tenure and raw frequencies cannot be discussed publicly or compared with other groups (http://sitemaker.umich.edu/lmt/terms). The measures are suitable for measuring growth in teacher knowledge, studying how teacher knowledge relates to student achievement, studying the mathematical knowledge that teachers need, and for looking at how that knowledge is organized (Hill et al. 2004).

Factor analyses of responses to the items can be compared to the hypothesized domains of teacher knowledge (Hill et al. 2004). Item response theory (IRT) models can be used to score the teachers' responses in standard deviations where the mean is 0 (Hill 2007). The difficulty of the items and their ability to discriminate among teachers are also estimated using IRT models. Studies were conducted in the United States to evaluate the validity of using the MKT measures. Although more work remains to be done on validation, and in understanding MKT and its measurement, even within the United States, Hill et al. (2007a) found that teachers' scores on the measures could predict mathematical features of the teachers' instruction and student achievement.

One early study by the group looked at how the MKT measures might be adapted for use outside the United States (Delaney et al. 2008). This study made four categories of changes to items: to the general cultural context, to the school cultural context, to the mathematical substance, and other changes. As other researchers have used the measures in various countries, including non-English speaking countries, they have proposed additional or alternative categories of changes to be considered (see Mosvold et al. 2009; and Kwon et al. 2012) when measures developed for use in one country are adapted for use elsewhere.

4 Core Results on the Professional Knowledge of Mathematics Teachers

4.1 Prospective Teachers

4.1.1 MCK and MPCK by Countries

Prospective primary teachers from Taiwan achieved the most favourable MCK result of all of the countries participating in the TEDS-M primary study (Blömeke et al. 2011b). The difference from the international mean was large—more than one standard deviation, which is according to Cohen (1988) a highly relevant difference. The achievement of primary teachers from the US was slightly above the international mean and roughly on the same level as the achievement of teachers in Germany and Norway. Their difference from the international mean was significant but of low practical relevance. These groups of teachers also reached significantly lower performance levels than Swiss and Thai teachers. If we take into account the Human Development Index used by the UN in order to indicate the social, economic, and educational developmental state of a country, the high performance of teachers from Russia and Thailand was striking.

Regarding MPCK, the achievement of prospective primary teachers from the US was roughly on the same level as the achievement of teachers in Norway, which was significantly above the international mean. In this case, the difference from the international mean was of practical relevance. Teachers from Singapore and Taiwan outperformed the US teachers. Whereas Singapore was behind Taiwan in the case of MCK, these countries were on the same level in the case of MPCK. Regarding MPCK, Norway and the US were only half of a standard deviation behind the two East Asian countries, whereas this difference reached one standard deviation regarding MCK.

Prospective lower secondary teachers from Taiwan, Russia, Singapore, Poland, and Switzerland significantly outperformed teachers from the other countries regarding MCK (Blömeke et al. 2012a). If we take into account the Human Development Index, the performance of lower secondary mathematics teachers from Russia and Poland was remarkable. Regarding MPCK, the achievement of Taiwanese and Russian teachers was outstanding. The achievement of teachers from Singapore, Switzerland, and Russia was also well above the international mean.

The ranking of countries in TEDS-M was very similar to the ranking of countries in TIMSS (Mullis et al. 2008), which allows a preliminary tentative conclusion that we are talking about a cyclic relationship—with the option to improve student achievement by increasing mathematics teachers' professional knowledge.

Hsieh, Lin, and Wang (this book) examined the TEDS-M data based on the notion of thought-oriented mathematical knowledge (Niss 2002). They found that although Taiwan outperformed Singapore, these two East Asian countries shared the same structural pattern in their responses (see also Hsieh et al. 2011). The pattern suggested that, when compared with other countries, Taiwanese and Singaporean teachers performed relatively better with respect to mathematical language, including representing mathematical entities and handling mathematical symbols or formalisms, than with respect to modeling and reasoning.

As another relative weakness of prospective teachers in Taiwan, the authors discovered that—when compared with other countries—Taiwan performed worse in diagnosing student achievement. This result confirmed findings of domestic studies in Taiwan that teachers demonstrated an incomplete understanding of student learning. Overall, these results suggest a specific cultural pattern in teacher performance in Taiwan as discussed in the literature (Leung 2001; Leung et al. 2006) and probably related to basic cultural features of the society (Hofstede 1983, 1993).

Senk, Tatto, Reckase, Rowley, Peck, and Bankov (this book) point out large structural variations across countries in how teachers were trained to teach mathematics. The authors group teacher training programs into four groups. Primary teachers trained as mathematics specialists tend to have higher MCK and MPCK than those trained as generalists. However, within each group of teacher training programs differences of about one to two standard deviations in MCK and MPCK occur between the highest and the lowest achieving countries. The authors infer from these results that the relative performance within countries may vary greatly, especially if more than one teacher training program exists.

MCK and MPCK are strongly correlated in most countries that took part in TEDS-M (Blömeke et al. 2010a, 2010b). However, the strength varies by country and level, which may be an indicator of cultural differences on the one hand and differences in the nature of primary and lower secondary teachers' knowledge on the other. With respect to primary teachers, the weakest correlation exists in Botswana, the Philippines, and Singapore (around $r = 0.30$) whereas it is strongest in Poland, Russia, and Germany (around $r = 0.60$). With respect to lower secondary teachers, the correlations are mostly stronger (around $r = 0.70$ in Germany, Russia, and Poland; around $r = 0.40$ in the Philippines, Switzerland, Oman, and Taiwan).[6]

[6] Botswana is an exception here with $r = 0.18$.

4.1.2 Factors Related to Prospective Mathematics Teachers' Professional Knowledge

According to the TEDS-M results, countries differed with respect to the demographic background of their prospective teachers and the opportunities to learn they received during teacher training (Tatto et al. 2012). Both the individual and the institutional characteristics levels influenced the acquisition of teachers' professional knowledge (Blömeke et al. 2012a; Schmidt, Houang, and Cogan, this book; Wong, Boey, Lim-Teo, and Dindyal, this book).

Influence of Individual Characteristics on Teacher Knowledge

Gender Effects With respect to prospective primary teachers from the 15 countries that participated in TEDS-M, we have to note significant achievement differences in favor of male compared with female teachers in most countries (Blömeke et al. 2011b). Pronounced gender gaps existed at the end of teacher training in particular with respect to MCK. However, the gender effect did not apply to the same extent to MPCK; in Malaysia, female teachers even outperformed male teachers. The MCK differences between male and female teachers were the largest in Poland, whereas only in a few countries (Malaysia, the Philippines, Botswana, and Germany) did no significant differences occur. The comparative study MT21 provided first evidence that gender-related achievement differences in MCK might apply to lower secondary teachers as well (Blömeke and Kaiser 2010).

The MCK gender effect may be cumulative. MCK has been built over a long period of schooling and teacher training, whereas MPCK was taught only at university. The well-known K-12 disadvantages in the mathematics achievement of girls compared with boys in Western countries (Hyde et al. 2008) may result in differences in prior knowledge as well as in motivational differences and thus play out in teachers' MCK.

That several countries were able to avoid gender inequalities indicates that teacher achievement may reflect cultural patterns. With respect to MPCK, its pedagogical nature may reduce women's disadvantages, too. Evidence exists that female future teachers tend to support pedagogical motives more strongly than male future teachers do, specifically in comparison to subject-specific motives (Eberle and Pollak 2006).

Language Effects Another background characteristic associated with prospective teachers' professional knowledge in some countries was their *language* background (Blömeke et al. 2011b). In Germany, the United States, and Thailand, differences of high practical relevance occurred in MCK as well as in MPCK. The differences were always in favor of those teachers whose first language matched the official language of instruction in teacher training. Thus, students with a minority background, speaking a different language at home from the language of instruction, were at a disadvantage.

This difference may result from selection effects during schooling. The language background is an important predictor of K-12 achievement (Coleman et al. 1966; Thomas and Collier 1997). Students with a different language background from the one used in instruction may have been filtered out during schooling or at the beginning of teacher training. An indicator for this interpretation is that the proportion of teachers with a different language background is lower in the prospective teacher force than on average in the K-12 student population. Again, several countries (e.g., Spain) were successful in avoiding differential language effects. It would be worthwhile to examine in detail how these countries accomplished language equity.

Prior Knowledge Characteristics strongly associated with prospective teachers' MCK and MPCK, not only in a few countries but more or less universally, were the perceived high-school achievement as well as the number of mathematics classes at school (Blömeke et al. 2012a). Effect sizes were large in both cases. Assuming that both predictors were appropriate to indicate prior knowledge, these results are in accordance with the general state of research (see, e.g., Anderson and Lebière 1998; Simmons 1995). Higher prior knowledge facilitates the acquisition of new knowledge, for example by supporting the integration of new information into existing schemata, the modification of knowledge structures, or the compilation and chunking of knowledge.

Motivation Effects A final set of individual characteristics associated with MCK and MPCK was motivation. Subject-related motives were positively related whereas extrinsic motives were negatively related to teacher training outcomes if other individual predictors were controlled (Blömeke et al. 2012a). It seems as if the persistence to overcome mathematics-related learning difficulties or to invest time and energy in the learning of mathematics decreases if somebody wants to become a teacher primarily because she wants the long-term security of the job but increases if she is interested in the subject (Wigfield and Eccles 2000).

Several countries concluded from these kinds of results that an active recruitment policy is necessary in order to ensure a high-quality pool of applicants from which teacher training can draw. This, however, raises questions of feasibility with respect to salary, prestige, and motivation as the teaching profession does not always compare favorably with other professions available to highly mathematically literate college graduates. In Singapore, all students selected into teacher training receive full salaries as if they were already practicing. Tuition fees are paid by the state (Wong, Boey, Lim-Teo, and Dindyal, this book). Similarly, South Korea and Taiwan recruit their future teachers from the best high school graduates. In contrast, top students in many Western countries choose law or medicine as study fields. Schmidt, Houang, and Cogan (this book) describe therefore with respect to the US a different measure to accomplish the same objective. If the US were to successfully implement the new federal mathematics curriculum, the achievement of students in eighth grade might look different and resemble more that of high-scoring countries. This might then result in better entrance characteristics even if future teachers were drawn from the same part of the distribution as is currently done. In any case, it is necessary to have in mind that teacher knowledge at the end of teacher training is partly an amalgam of the teachers' educational background and cultural and social characteristics.

Institutional Characteristics Prior to TEDS-M, the state of research had indicated large cultural diversity in the curricula of teacher training across countries (Bishop 1988). However, a study by Adler et al. (2005) analyzed 160 papers about mathematics teacher training and pointed out that most studies were restricted to small-scale qualitative research and English-speaking countries. In quantitative studies, only the type of license or the number of courses taken was used to define OTL. These measures reflected the amount of content coverage without taking into account which content was offered.

TEDS-M was the first study that provided in-depth information about OTL. Blömeke and Kaiser (this book) summarize with respect to primary teacher training that, in fact, a comprehensive core curriculum accepted in all TEDS-M countries neither existed in mathematics pedagogy nor in mathematics. At the same time, the authors were able to conclude that the heterogeneity may be less pronounced than usually discussed. In mathematics (number theory and probability) and mathematics pedagogy (teaching methods), topics existed which were taken by most prospective primary teachers. Blömeke (2013) reached a similar conclusion with respect to prospective lower-secondary teachers.

Furthermore, it was sufficient to distinguish between a few profiles of OTL in mathematics ("advanced university mathematics", "basic university mathematics", and a restriction to "school mathematics") as well as in mathematics pedagogy (a "broad mathematics pedagogy curriculum", a "functional mathematics pedagogy curriculum", and "teaching methods") to describe appropriately the mathematics and mathematics pedagogy curriculum of primary teacher training across the participating countries. The dominating philosophy across the TEDS-M countries was to provide OTL in basic university mathematics and a broad mathematics pedagogy curriculum. A recent study of Blömeke (2013) on lower-secondary teacher education revealed a similar pattern.

The common topics and the low number of profiles may reflect shared visions of what primary and lower-secondary teachers are supposed to know before they enter the profession. This result confirmed a prior result from a comparative study on lower-secondary mathematics teacher training in six countries (Blömeke et al. 2008b; Schmidt et al. 2011). Its data indicated that specific OTL profiles may exist and that these may have been influenced by culture: in five countries, the multiple institutions where teacher training took place tended to cluster together with respect to the OTL offered, suggesting country-level agreement reflecting a cultural effect (Schmidt et al. 2008).

However, a closer examination of the OTL also revealed patterns of high-achieving vs. low-achieving TEDS-M countries. The level of mathematics and the emphasis of mathematics teaching practice were higher in the first cases (Hsieh et al. 2013). Schmidt, Houang, and Cogan (this book) confirmed these aggregated country-level results through multi-level modeling. They found within-country evidence of a significant relationship between teacher training and both MCK and

MPCK. OTL in mathematics and mathematics pedagogy predicted an increase in MCK of about half of a standard deviation, which is a substantial effect size. The largest single effect was produced by OTL closely related to experiences with mathematics instruction. OTL in mathematics had not only a strong direct influence on MCK, though, but also on MPCK, and they probably mediated the effects of OTL in mathematics pedagogy (Blömeke et al. 2012a).

Wong, Boey, Lim-Teo, and Dindyal (this book) summarized in this sense the evidence for Singapore that teacher training counts. Even well-qualified recruits require proper training to realize their potential. Therefore, programs are regularly revised in Singapore to ensure that they are responsive to both external changes such as recruitment numbers and education initiatives launched by the state and within-institution research and feedback from the student teachers about their training.

The TEDS-M results are in line with recent evidence from large national longitudinal studies. These had revealed that pure structural features, such as program or degree type, do not have significant effects on short-term outcomes of teacher training, such as teacher knowledge, or long-term outcomes, such as teacher retention or student achievement (Goldhaber and Liddle 2011). In contrast, especially in the case of mathematics teachers, the evidence increasingly suggests that the quality of programs does have an impact on teacher outcomes (Boyd et al. 2009; Constantine et al. 2009). Content courses in mathematics are an important part of these quality features as they provide the background knowledge and the conceptual and factual knowledge necessary to present mathematics topics to learners in a meaningful way and to connect the topics to one another as well as to the learner's prior knowledge and future learning objectives (Cochran-Smith and Zeichner 2005; Wilson et al. 2001). In this context, the curriculum sequence and delivery seems to have an important influence on graduates' subject matter knowledge (e.g., Tatto et al. 2010).

Knowing the content, however, provides only a foundation for mathematics teaching. Student achievement is higher if a strong subject-matter background is combined with strong educational credentials (Clotfelter et al. 2006). The importance of professional preparation, specifically the understanding of how learners acquire mathematical knowledge, how to teach racially, ethnically, and linguistically diverse students, and using a wide array of instructional strategies, represents another robust finding of teacher education research across various studies (Constantine et al. 2009; NRC 2003). A third robust finding on the impact of OTL on the outcomes of teacher education is the quality of the teaching methods experienced—in particular, the opportunity to engage in actual teaching practices, such as planning a lesson or analyzing student work, rather than only listening to lectures (Boyd et al. 2009).

4.2 Practicing Mathematics Teachers

MKT is a practice-based theory in the sense that MKT is identified with reference to the practice of teaching. Because the practice of teaching can vary from country to country, the content of MKT could vary accordingly. Furthermore, the MKT measures differ from other instruments for measuring teacher knowledge because they are not criterion referenced (Hill et al. 2007b). Criteria stating what mathematical knowledge teachers should hold have not been specified. Instead the measures were designed so that teachers could be ordered "relative to one another and to the underlying trait being assessed" (p. 131). Therefore the studies of the MKT held by practicing teachers in various countries could not be compared as the knowledge of prospective teachers could be in the TEDS-M study. In some cases, the researchers' goal was not so much to measure teachers' knowledge at present but to develop guidelines to enable such study in the future.

For example, the study by Kwon et al. (2012) considers changes made to the multiple-choice items in order to use them with Korean teachers. Despite initially proposing a more elaborate scheme of potential changes to items and despite making various changes, such as replacing terminology relating to base ten materials with an alternative term and subsequently with a diagram, the authors conclude by advocating a conservative approach to adaptation in order to maintain the integrity of the items. They identify potential risks that are inherent in adaptation, such as reducing or increasing the mathematical demand of the items or losing the validity of an item. Such an argument is helpful for others who seek to adapt the items for use outside the United States.

The study by Fauskanger et al. (2012) describes an iterative process that can be used to determine how well items—adapted or not—work when administered to a group of teachers outside the United States. Point biserial correlations (how teachers' performances on a given item correlate with their performances on all other items) and the relative ordering of item difficulties (measured in standard deviations where an item that a teacher with average ability has a 0.5 chance of responding correctly has a difficulty of 0) for US and Norwegian teachers was studied in order to identify items which could be probed further in focus group interviews. Although different causes of the problems were hypothesized and the precise problems ultimately remain somewhat uncertain, possible reasons include differences in defining mathematical objects, the use of unfamiliar teaching contexts, and problematic aspects of translation. When combined with the approach of Kwon and her colleagues (2012), the findings of the Norwegian study provide a means of identifying problematic items; but mindful of the risks inherent in any adaptation, changes should only be made if a compelling reason exists to do so.

Ng (2012) studied Indonesian teachers' performances on geometry items. Having studied the point biserial correlations, the item difficulties and the reliability of the measures and compared them to similar statistics among US teachers, he concludes that the geometry measures "may not be a good set of assessments to evaluate Indonesian teachers' mathematical knowledge for teaching geometry". He believes that a reason for this is attributable to national differences between the United States and Indonesia with regard to differences in how shapes are classified. Nevertheless, he believes that the process of studying the performance outside the United States of measures based on MKT can contribute to the overall development of our understanding of MKT.

Cole (2012) and Delaney (2012) evaluate the validity of using the MKT measures in Ghana and Ireland respectively. Cole looks at the consistency between three teachers' scores on individual items and the mathematical reasoning for choosing the responses they chose. Cole finds that most items could be used validly in Ghana despite evidence of "cultural incongruence". Sources for the incongruence came from possible differences in teaching practices, a specific question format, the length of time taken to do the test, and taking the test in English where despite English being the language of schooling, none of the teachers in the study were native English speakers.

Delaney (2012) attempted to validate the use of the measures to study Irish teachers' MKT. Like Cole (2012), he found that, in general, teachers' thinking was consistent with their responses to the measures. The factors found among teachers' responses to the items in Ireland are similar to the factors found among US teachers' responses. However, the organization of the factors in both countries differs from the hypothesized domains. Finally, Delaney (2012) found that only in the case of five of ten teachers did their MKT score predict the mathematical quality of instruction to be found in their teaching. This contrasted with findings in a similar study in the United States (Hill et al. 2007a). He concludes by identifying challenges to validating the use of the MKT measures in settings outside the United States.

Small-scale comparative studies of practicing teachers have been conducted using measures and frameworks other than MKT. An et al. (2004) compared teachers' PCK between Chinese and US groups, focused on fractions, ratio, and proportion. Based on Shulman's work, they conceptualized pedagogical content knowledge as an amalgam of content and pedagogical knowledge. As Ma (1999) also found, US teachers performed more poorly than Chinese teachers on the PCK measure. Further comparisons of opportunities to learn PCK in China and the US revealed that the Chinese had gained much of their knowledge through school-based pre-service training led by nationally well-known expert teachers and continuous professional development activities, especially by observing each other's lessons and jointly discussing them (An et al. 2004; Paine 1997; Paine and Ma 1993).

In more recent work, An (2009) studied 385 elementary school teachers at 37 schools in six cities/regions in four provinces in China. The association between CK and PK in three areas of content and six areas of pedagogy in multi-digit division was medium (Cramer's $V = 0.33$). The validation findings revealed that Chinese teachers' pedagogical knowledge was related to grade levels taught and to the majors they had chosen in college. These results provide evidence that their instruments in fact are able to measure meaningful differences between CK and PK in multi-digit division, but An concludes that further studies are needed to test PCK in other content areas.

5 Challenges of Comparative Large-Scale Assessments of Teacher Knowledge

It is a methodological challenge to assess teacher knowledge from a comparative perspective. Research perspectives have to be adjusted across borders and educational traditions. Furthermore, it is a methodological challenge to assess the development of knowledge among prospective teachers in the context of a differentiated tertiary education system. Not only do a variety of institutions, teacher training programs, and job requirements exist, but also the outcome is hard to define and even harder to measure.

5.1 Generalizability

In most of the studies reported above, a definition of teacher knowledge as a context-specific disposition was applied. The knowledge can be acquired and it is needed to cope successfully with domain-specific classroom situations and tasks. However, several controversies are unsolved. What is the role of attitudes and beliefs in this context? In future research, in addition to MCK and MPCK as subject-specific facets of mathematics teachers' knowledge, other cognitive criteria like general pedagogical knowledge or affective characteristics like teacher beliefs should be included (as was done in TEDS-M) in order to develop a full model. Such an approach would increase the validity of studies on teacher knowledge.

With respect to theory development and generalizability, it seems important not to focus solely on analytical approaches as used in TEDS-M and LMT. Shavelson (2012) unpacks competency as a complex ability construct closely related to real-life situation performance. He exemplifies how to make it amenable to measurement in a holistic way by research from business, military, and education sectors. The generalizability theory, a statistical theory for modeling and evaluating the dependability of competency scores, is applied to several of these examples. The paper then pulls the threads together into a general competency measurement model.

Shavelson (2012) points out that there are limitations to measuring competency in terms of resources, costs, and time on various levels. Performance assessment is an issue that has long been discussed (Kane 1992). It is difficult to generalize results from one situation to another, that is, there are problems with reliability (Brennan and Johnson 1995), and it is difficult to validate the measures (Kane et al. 1999). How representative are, for example, the situations to be worked on in a measurement?

We have presented the TEDS-M framework in some detail (see Table 1). Although it looks convincing, a comparison with the way that California evaluates its pre-service mathematics teachers' knowledge (Wu 2010) reveals that different approaches can be taken. The Teacher Performance Assessment (TPA) depicts classroom situations in a standardized way and according to the state's standards, "Teaching Performance Expectations". Four tasks have to be dealt with: connecting instructional planning to student characteristics, assessment, lesson design, and reflection. These have to be applied to (only) two groups of learners which are not present in the TEDS-M framework: English language learners and special education students. The difference between the TEDS-M and the Californian approaches reveals very different visions of what mathematics teachers are supposed to know and be able to do.

It is worthwhile to examine the phenomenon of context dependability and generalizability in more detail—also with new instruments beyond the limitations of paper-and-pencil tests. Approaches in this respect exist in several countries. Wu and Li (2008), for example, examine "the power to perform a teaching task" as the criterion of applying knowledge in the classroom. They videotaped 119 lessons from nine K-8 teachers, interviewed them, and assessed student achievement of these teachers. Observation instruments were designed based on components of the PCK (An et al. 2004). They discovered distinct patterns in teaching performance with differences in their relationship to student achievement. The TIMSS and TIMSS-R video studies had provided the first comparative evidence in this direction (though without testing teacher knowledge).

In this context, the slight difference in spelling between "competency" and "competence" in the English language may be relevant for evaluating the generalizability of results. Sadler (2012) points out that a conceptual distinction can be made between the two terms, which in turn leads to distinct measurement approaches. A "competency" often means an identifiable practice. "Competence", in contrast, often consists of a large number of discrete competencies—as is assumed in TEDS-M—which could be tested independently. Competence involves being able to select from and then orchestrate a set of competencies to achieve a particular end within a particular context. The competent person makes multi-criterion judgments that are consistently appropriate and situation-sensitive. What is more, the range of situations faced by many professional practitioners is potentially infinite. Decomposing competence into manageable components in order to facilitate judgments has value in certain contexts, but the act of decomposition can obscure how a practitioner would juggle the various bits together to form a coherent whole. It may be worthwhile to follow this assumption with research projects that compare the results of more integrative and holistic judgments with the results of more analytic approaches.

Also, it is necessary to consider the knowledge development of prospective and practicing teachers along a continuum of lifelong learning (Musset 2010). Such an approach would also allow including classroom observations of teacher performance and possibly even student achievement in order to examine the construct validity of measures.

5.2 Benefits and Limits of Comparative Research

Like everyone else, researchers are embedded in their own culture, and so they often overlook matters of culture. This is particularly the case for teacher training, given the unique way in which it incorporates or touches upon many different levels of education and stands at the intersection of education and other social, economic, and political forces (Blömeke and Paine 2008). This embedded character of the system of teacher training in any one country makes looking beyond that country's experience mandatory in order to recognize the assumptions which drive it, which are all too often taken for granted. The investigation of another teacher training system in a foreign country, for example, and the discovery that it is possible to organize the training differently, sheds new light on domestic systems. The recognition of this cultural boundedness of teacher training is an argument for approaching a comparative study in ways that maximize opportunities for cross-cultural communication and the direct examination of concepts (LeTendre 1999).

As such, language problems become important in comparative studies and are far more demanding to resolve than a "simple" translation of instruments or responses (NRC 2003). Of course, at one level, this is a common, familiar and well-studied aspect of cross-cultural studies, for which there are now widely-used conventions of translation, back translation and so on (Hambleton 2002). In teacher training, however, more language-related challenges exist that require attention. They are a problem of cultural boundaries. In some countries, Ghana for instance, the language of schooling may vary from the language of the home for many students. Many terms from native languages cannot be translated because adequate English terms are missing and vice versa. It is even difficult to name the process by which future teachers learn their profession: is it teacher education, is it teacher training, or is it perhaps teacher preparation?

These questions relate to deeper and often tacit assumptions about schooling, teaching, and learning to teach. As these terms connect to broadly shared cultural beliefs, the uniqueness of their meaning often is not explicit and can easily escape scrutiny unless outsiders to the cultural community stumble over them and begin to enquire about them (Blömeke and Paine 2008). Behind the apparently simple choice of whether to refer to the practice as teacher education, teacher training, teacher preparation, or something else, lie other aspects of history, policy, social values, and cultural norms. These are worth examining in detail.

These conceptual challenges of comparative research are extended by methodological challenges. Owing to the low number of countries, in the TEDS-M analyses a "one size fits all approach" (van Ewijk and Sleegers 2010) has to be used, though with parameter estimates the same for all countries in multi-level analyses (see, for example, Blömeke et al. 2012a). Thus, a risk exists that country-specific variation in the effect sizes of some predictors is overlooked. At least for the larger countries in the TEDS-M sample, it seems therefore worthwhile to estimate country-specific models in addition to cross-country models.

Future research should also examine in more detail the question of measurement invariance in TEDS-M (Blömeke et al. 2011a). Van de Vijver (1998, p. 43) points to a serious threat: "An instrument is biased if its scores do not have the same psychological meaning across the cultural groups involved; more precisely, an instrument is biased if statements about (similarities and differences of) its scores do not apply in the psychological domain of the scores." Equivalence is thus the objective to be achieved. It consists of several dimensions (Vandenberg and Lance 2000): conceptual equivalence of the latent trait in each group, equivalent associations between operationalizations in each group, and the extent to which they are influenced to the same degree by the same factors. Despite agreement on the importance of achieving equivalence, however, many terms have been used to describe aspects of equivalence. Johnson (1998) has identified over fifty of them, several of which overlap and many of which have not been well defined.

Hierarchical IRT and multiple-group confirmatory factor analysis provide the tools to determine important properties such as configural invariance, metric invariance, and scalar invariance (Fox 2005; Vandenberg and Lance 2000). Even if full invariance—which is rarely accomplished in cross-cultural research—cannot be determined for TEDS-M, such studies would reveal the extent to which partial invariance is supported. Approaches could then be taken to deal appropriately with the problems. Using hierarchical IRT, for example, De Jong et al. (2007) were able to relax all invariance requirements across groups while retaining the possibility of making substantive comparisons. Such studies would be relevant not only with respect to the TEDS-M assessment data but also, and perhaps more importantly, with respect to the OTL and beliefs data, given the likelihood of self-reported data being even more vulnerable to bias (Blömeke et al. 2010a, 2010b).

5.3 Research on Teacher Knowledge as a Tool to Improve Teacher Training

Wong, Boey, Lim-Teo, and Dindyal (this book) discuss the value of studies such as TEDS-M for teacher training. They point out that the released MCK and MPCK items can be used as a training resource. In fact, the Singaporean TEDS-M team is preparing a book consisting of these released items, the scoring guides, the Singapore results against international benchmarks, and samples of constructed responses. Teacher educators can then use these materials in their lessons with future cohorts of prospective teachers by, for example, exploring strategies to remedy misconceptions, designing classroom activities that mirror the scenarios described in the TEDS-M items, and linking the assessment items to the TEDS-M framework and thus analyze conceptions of teacher knowledge. Thus, although the TEDS-M items were originally created as a summative assessment of teacher knowledge at the end of teacher training, they can be used as a formative assessment of teacher knowledge.

Teacher educators may also want to compare the outcomes of different programs and different institutions in their country. Within almost all countries, huge between-program disparity existed. This means that within the same cultural context some institutions are doing better than others. They may represent a benchmark and provide important information about features of teacher education which can be more easily adapted than features from other countries. In particular, the structure and content of the mathematics and the mathematics pedagogy curriculum should be put to the test.

From those countries achieving high scores in TEDS-M, we may want to learn about promising ideas on how to organize teacher training programs. Again the Singapore example may serve as a role model (Wong, Boey, Lim-Teo, and Dindyal, this book). Here, mathematicians teach the content courses and mathematics educators teach the pedagogical content courses, but they belong to the same department. Under this organization, there are many opportunities for mathematicians and mathematics educators to work in committees and projects that draw on their separate expertise to achieve the common goal of training competent mathematics teachers. They can also share information about the prospective teachers. Furthermore, all mathematicians learn to supervise practica of prospective teachers at secondary schools through a process of mentoring, and this requirement provides an important opportunity for them to observe first-hand school mathematics teaching and to share their views as a subject specialist with the prospective teachers.

For achieving an increase of teacher education effectiveness, the TEDS-M study points to two further measures, each with separate effects. Providing OTL in mathematics as well as increasing entrance selectivity may have positive consequences for the outcomes of teacher training and thus in the long run for student achievement in mathematics. Mathematics is one of the most important school subjects and a gatekeeper to academic and professional success. Investments in the training of teachers should therefore pay off quickly. Entrance selectivity is a sensitive issue, however. Not everywhere is teaching such a popular and rewarding job that enough applicants for teacher education are available. Higher selectivity, however, may increase the reputation of the profession in the long run so that institutions can recruit from a larger pool.

In addition to such reforms, policymakers have to be aware of the continuing problem of societal inequalities in teacher education outcomes. Special support of female teachers when it comes to the acquisition of MCK in order to overcome cumulative disadvantages of a long history of K-12 schooling seems to be a meaningful measure in many TEDS-M countries.

5.4 Adaptation

Because MKT is a practice-based theory (Ball and Bass 2003), and teaching practices may be cultural in nature (Stigler and Hiebert 1999), differences may exist in relation to the form MKT takes in different countries. In order that teachers in a given country can respond to the items without being distracted by names or contexts that would be unusual or non-existent for them, some adaptation is necessary. Indeed, all of the studies of MKT adapted the measures they used in their research. However, such adaptations must avoid distorting the mathematical content or the mathematical demands of the measures. Guidelines for adaptation exist for studies such as TIMSS (Johansone and Malak 2008) and PISA, and similar guidelines will be needed for measures of MKT if the measures are to be used in diverse countries around the world. Recommendations from the studies in the 2012 ZDM special issue (see, e.g., the papers by Cole 2012; Kwon et al. 2012) could be useful in developing such guidelines.

5.5 Validation

Despite the challenges found by Cole (2012) and Delaney (2012), and despite the general absence of validity analyses in educational assessments (Hill et al. 2007b), validation of the use of the measures needs to be an important part of studies of MKT in any countries in which the measures are used. Attending to this will help to clarify the organization of the subdomains of MKT. It will also ensure that the items are tapping knowledge that is needed to teach mathematics. Above all, it will ensure that the measures are valid for the uses to which they will be put.

6 Concluding Remarks

This review has presented an overview of research on the assessment of mathematics teachers' knowledge as one of the most important parameters of school quality. Its focus has been on comparative and international studies that allow for analyzing the cultural dimensions of teacher knowledge. We presented in detail the conceptual frameworks underlying TEDS-M and MKT/LMT, the instruments designed to assess the content knowledge and pedagogical content knowledge of future and practicing mathematics teachers in different countries, and core results of its level and structure.

Although cross-national and comparative surveys of student knowledge have a longer track record and attract more sustained funding and attention than studies of teacher knowledge, the current book acknowledges the importance of learning about teacher knowledge by studying it beyond individual countries. More needs to be learned about adapting and validating measures for use in other countries. However, studying across countries has the potential to offer insights into the original frameworks and contribute to a better and clearer conception of the frameworks underlying TEDS-M and MKT/LMT.

It appears as if—not only in the context of TEDS-M and MKT/LMT—the research on mathematics teacher knowledge has made important progress. Shulman's model of teacher knowledge leads many studies so that the traits examined intend to represent the same. Our summary in Sect. 1 demonstrates this intention with respect to TEDS-M and LMT. The similarities in turn make it easier to compare the instruments and to connect the studies' results to each other than was the case in prior research.

Still, many challenges exist. Cross-country equivalence of meaning and predictive validity are the most important ones. The many studies connected to LMT had their focus on this perspective. LMT started as a one-nation enterprise but is expanding to many countries all over the world. In contrast, TEDS-M started as a comparative study but is now complemented by many national studies that go into more detail. We are starting to get ideas about how teacher knowledge develops and how it is connected to teacher education and student achievement. It turns out, as assumed, that teacher knowledge is the crucial link between mathematics teacher education and student achievement in mathematics. How much, and in what quality, opportunities to learn are provided significantly influences the knowledge achieved during teacher training.

In turn, teacher knowledge represents an important predictor of student achievement because a mathematics teacher's decision making in class is a function, among others, of her mathematical knowledge (Schoenfeld 2010). Thus, the two perspectives together—research on prospective and practicing teachers—provide an appropriate view on the relationship between mathematics teacher education and what is accomplished in schools, although multiply mediated by complex context factors. Here, much further research is needed so that we will be able to understand the nature of teacher competencies underlying classroom performance.

References

Adams, R. J., Wilson, M., & Wang, W. C. (1997). The multidimensional random coefficients multinominal logit. *Applied Psychological Measurement*, *21*, 1–24.

Adler, J., & Ball, D. (2009). Introduction to and overview of this special issue. *For the Learning of Mathematics: Knowing and Using Mathematics in Teaching*, *29*(3), 2–3.

Adler, J., & Davis, Z. (2006). Opening another black box: researching mathematics for teaching in mathematics teacher education. *Journal for Research in Mathematics Education*, *37*(4), 270–296.

Adler, J., Ball, D., Krainer, K., Lin, F. L., & Novotna, J. (2005). Reflections on an emerging field: research on mathematics teacher education. *Educational Studies in Mathematics*, *60*, 359–381.

An, S. (2009). Chinese teachers' knowledge of teaching multi-digit division. *Journal of Mathematics Education*, *2*(1), 27–54.

An, S., & Wu, Z. (2011). Enhancing mathematics teachers' knowledge of students' thinking from assessing and analyzing misconceptions. *International Journal of Science and Mathematics Education*, 1–37.

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.

Anderson, J. R., & Lebière, C. (1998). *The atomic components of thought*. Hillsdale: Lawrence Erlbaum.

Ball, D. L., & Bass, H. (2003). *Toward a practice-based theory of mathematical knowledge for teaching*. Paper presented at the Annual meeting of the Canadian Mathematics Education Study Group, Edmonton, AB.

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.

Bateson, M. C. (1994). *Peripheral visions: learning along the way*. New York: Harper Collins.

Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Stanat, P., Tillmann, K.-J., & Weiß, M. (Eds.) (2001). *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich.

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.

Bednarz, N., & Proulx, J. (2009). Knowing and using mathematics in teaching: conceptual and epistemological clarifications. *For the Learning of Mathematics: Knowing and Using Mathematics in Teaching*, *29*(3), 11–17.

Begle, E. G. (1972). *Teacher knowledge and student achievement in algebra* (School Mathematics Study Group Reports, No. 9). Washington, D.C.: National Science Foundation.

Begle, E. G. (1979). *Critical variables in mathematics education*. Washington: Mathematical Association of America and the National Council of Teachers of Mathematics.

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.

Bishop, A. (1988). *Mathematical enculturation: a cultural perspective on mathematics education*. Dordrecht: Kluwer.

Blömeke, S. (2013). Content, professional preparation and teaching methods: how diverse is teacher education across countries? *Comparative Education Review*.

Blömeke, S., & Kaiser, G. (2010). Mathematics teacher education and gender effects. In H. Forgasz, K.-H. Lee, J. R. Becker, & O. B. Steinstorsdottir (Eds.), *International perspectives on gender and mathematics education* (pp. 263–283). Charlotte: Information Age Publishing.

Blömeke, S., & Paine, L. (2008). Getting the fish out of the water: considering benefits and problems of doing research on teacher education at an international level. *Teaching and Teacher Education*, *24*(4), 2027–2037.

Blömeke, S., Felbrich, A., Müller, Ch., Kaiser, G., & Lehmann, R. (2008a). Effectiveness of teacher education. State of research, measurement issues and consequences for future studies. *ZDM—The International Journal on Mathematics Education*, *40*(5), 719–734.

Blömeke, S., Kaiser, G., & Lehmann, R. (Eds.) (2008b). *Professionelle Kompetenz angehender Lehrerinnen und Lehrer. Wissen, Überzeugungen und Lerngelegenheiten deutscher Mathematikstudierender und -referendare – Erste Ergebnisse zur Wirksamkeit der Lehrerausbildung*. Münster: Waxmann.

Blömeke, S., Kaiser, G., & Lehmann, R. (Eds.) (2010a). *TEDS-M 2008: Professionelle Kompetenz und Lerngelegenheiten angehender Primarstufenlehrkräfte im internationalen Vergleich [Cross-national comparison of the professional competency of and learning opportunities for future primary school teachers]*. Münster: Waxmann.

Blömeke, S., Kaiser, G., & Lehmann, R. (Eds.) (2010b). *TEDS-M 2008: Professionelle Kompetenz und Lerngelegenheiten angehender Mathematiklehrkräfte für die Sekundarstufe I im internationalen Vergleich [Cross-national comparison of the professional competency of and learning opportunities for future secondary school teachers of mathematics]*. Münster: Waxmann.

Blömeke, S., Houang, R., & Suhl, U. (2011a). *IERI monograph series: issues and methodologies in large-scale assessments: Vol. 4*. *TEDS-M: diagnosing teacher knowledge by applying multidimensional item response theory and multi-group models* (pp. 109–126).

Blömeke, S., Suhl, U., & Kaiser, G. (2011b). Teacher education effectiveness: quality and equity of future primary teachers' mathematics and mathematics pedagogical content knowledge. *Journal of Teacher Education*, *62*(2), 154–171.

Blömeke, S., Suhl, U., Kaiser, G., & Döhrmann, M. (2012a). Family background, entry selectivity and opportunities to learn: what matters in primary teacher education? An international comparison of fifteen countries. *Teaching and Teacher Education*, *28*, 44–55.

Blömeke, S., Suhl, U., & Döhrmann, M. (2012b). Zusammenfügen was zusammengehört. Kompetenzprofile am Ende der Lehrerausbildung im internationalen Vergleich. *Zeitschrift für Pädagogik*, *58*(4), 422–440.

Blumer, H. (1969). *Symbolic interactionism: perspective and method*. Berkeley: University of California Press.

Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: a perspective on assessment and intervention. *Applied Psychology: An International Review*, *54*(2), 199–231.

Boyd, D., Grossman, P. L., Lankford, H., Loeb, S., & Wyckoff, J. (2009). Teacher preparation and student achievement. *Educational Evaluation and Policy Analysis*, *31*(4), 416–440.

Bramald, R., Hardman, F., & Leat, D. (1995). Initial teacher trainees and their views of teaching and learning. *Teaching and Teacher Education*, *1*, 23–31.

Brennan, R. L., & Johnson, E. G. (1995). Generalizability of performance assessments. *Educational Measurement: Issues and Practice*, *14*(4), 9–12.

Bromme, R. (1994). Beyond subject matter: a psychological topology of teachers' professional knowledge. In R. Biehler, R. W. Scholz, R. Straesser, & B. Winkelmann (Eds.), *Mathematics*

didactics as a scientific discipline: the state of the art (pp. 73–88). Dordrecht: Kluwer.
Bromme, R. (1997). Kompetenzen, Funktionen und unterrichtliches Handeln des Lehrers [Competencies, objectives and classroom performance of teachers]. In F. E. Weinert (Ed.), *Enzyklopädie der Psychologie: Psychologie des Unterrichts und der Schule. Vol. 3 [Encyclopedia psychology: psychology of teaching, learning and schools]* (pp. 177–212). Göttingen: Hogrefe.
Brouwer, N. (2010). Determining long term effects of teacher education. In P. Peterson, E. Baker, & B. McGaw (Eds.), *International encyclopedia of education* (Vol. 7, pp. 503–510). Oxford: Elsevier.
Brown, D. F., & Rose, T. J. (1995). Self-reported classroom impact of teachers' theories about learning and obstacles to implementation. *Action in Teacher Education*, *17*(1), 20–29.
Brown, S., & McIntyre, D. (1983). Teacher attitudes and innovation characteristics. Let's try again: a rejoinder. *Curriculum Inquiry*, *13*, 447–451.
Burgess, T. (2009). Statistical knowledge for teaching: exploring it in the classroom. *For the Learning of Mathematics*, *29*(3), 18–21.
Burghes, D. (2008). *International comparative study in mathematics teacher training*. Reading: CfBT.
Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: a theoretical synthesis. *Review of Educational Research*, *65*(3), 245–281.
Butterfield, B., & Chinnappan, M. (2010). Walking the talk: translation of mathematical content knowledge to practice. In L. Sparrow, B. Kissane, & C. Hurst (Eds.), *Proceedings of the 33rd annual conference of the mathematics education research group of Australasia: shaping the future of mathematics education* (Vol. 2, pp. 109–116). Freemantle: MERGA Inc.
Butterfield, B., & Chinnappan, M. (2011). Teacher knowledge activated in the context of designing problems. In J. Clark, B. Kissane, J. Mousley, T. Spencer, & S. Thornton (Eds.), *Mathematics: traditions and (new) practices. Proceedings of the 23rd biennial conference of the Australian Association of Mathematics Teachers Inc. and the 34th annual conference of the Mathematics Education Research Group of Australasia Inc.* (pp. 142–150). Adelaide: MERGA/AAMT.
Cai, J. (2000). Mathematical thinking involved in U.S. and Chinese students' solving of process-constrained and process-open problems. *Mathematical Thinking and Learning*, *2*(4), 309–340.
Calderhead, J. (1996). Teachers: beliefs and knowledge. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 709–725). New York: Macmillan.
Chick, H. L., Pham, T., & Baker, M. K. (2006). Probing teachers' pedagogical content knowledge: lessons from the case of the subtraction algorithm. In P. Grootenboer, R. Zevenbergen, & M. Chinnappan (Eds.), *Identities, cultures and learning spaces. Proceedings of the 29th annual conference of the mathematics education research group of Australasia* (pp. 139–146). Canberra: MERGA.
Chinnappan, M., & Lawson, M. J. (2005). A framework for analysis of teachers' geometric content knowledge and geometric knowledge for teaching. *Journal of Mathematics Teacher Education*, *8*, 197–221.
Clotfelter, C., Ladd, H., & Vigdor, J. (2006). *Teacher credentials and student achievement in high school: a cross-subject analysis with student fixed effects*. Cambridge: National Bureau of Economic Research.
Cochran-Smith, M. & Zeichner, K. M. (Eds.) (2005). *Studying teacher education. The report of the AERA panel on research and teacher education*. Mahwah: Lawrence Erlbaum.
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale: Lawrence Erlbaum.
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.
Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, F., Mood, A. M., Weinfeld, F. D., et al. (1966). *Equality of educational opportunity*. Washington: U.S. Government Printing Office.
Constantine, J., Player, D., Silva, T., Hallgren, K., Grider, M., Deke, J., & Warner, E. (2009). *An evaluation of teachers trained through different routes to certification* (NCEE 2009-4043). Washington, DC: IES.

Cuoco, A. (2001). Mathematics for teaching. *Notices of the American Mathematical Society*, *48*(2), 168–174.

Darling-Hammond, L. (2000). Teacher quality and student achievement. *Education Policy Analysis Archives*, *8*, 1. http://epaa.asu.edu/ojs/article/view/392. Accessed on 3 May 2012.

De Jong, M. G., Steenkamp, J. B. E. M., & Fox, J.-P. (2007). Relaxing cross-national measurement invariance using a hierarchical IRT model. *Journal of Consumer Research*, *34*, 260–278.

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. 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.

Eberle, T., & Pollak, G. (2006). Studien- und Berufswahlmotivation von Passauer Lehramtsstudierenden. In N. Seibert (Ed.), *Paradigma. Beiträge aus Forschung und Lehre aus dem Zentrum für Lehrerbildung und Fachdidaktik* (pp. 19–37). Passau: Zentrum für Lehrerbildung und Fachdidaktik.

Eisenberg, T. A. (1977). Begle revisited: teacher knowledge and student achievement in algebra. *Journal for Research in Mathematics Education*, *8*(3), 216–222.

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.

Ferrini-Mundy, J., Floden, R., McCrory, R., Burrill, G., & Sandow, D. (2005). *A conceptual framework for knowledge for teaching school algebra*. East Lansing: Authors.

Forrester, T., & Chinnappan, M. (2011). Two avatars of teachers' content knowledge of mathematics. In J. Clark, B. Kissane, J. Mousley, T. Spencer, & S. Thornton (Eds.), *Mathematics: traditions and (new) practices. Proceedings of the 23rd biennial conference of the Australian Association of Mathematics Teachers Inc. and the 34th annual conference of the Mathematics Education Research Group of Australasia Inc.* (pp. 261–269). Adelaide: MERGA/AAMT.

Fox, J.-P. (2005). Multilevel IRT using dichotomous and polytomous items. *British Journal of Mathematical and Statistical Psychology*, *58*, 145–172.

Goldhaber, D., & Liddle, S. (2011). *The gateway to the profession. Assessing teacher preparation programs based on student achievement* (CEDR Working Paper 2011-2). Seattle, WA: University of Washington.

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. The international handbook of mathematics teacher education* (Vol. 1, pp. 117–132). Rotterdam: Sense Publisher.

Hambleton, R. K. (2002). Adapting achievement tests into multiple languages for international assessments. In A. C. Porter & A. Gamoran (Eds.), *Methodological advances in cross-national surveys of educational achievement* (pp. 58–79). Washington: National Academy Press.

Heid, K. M. (2008). *Mathematical knowledge for secondary school mathematics teaching*. Paper presented at the 11th International Congress on Mathematical Education, Monterrey, Mexico.

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. (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., Schilling, S. G., & Ball, D. L. (2004). Developing measures of teachers' knowledge for teaching. *The Elementary School Journal*, *105*(1), 11–30.

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., Blunk, M. L., Goffney, I. M., & Rowan, B. (2007a). Validating the ecological assumption: the relationship of measure scores to classroom teaching and student learning. *Measurement: Interdisciplinary Research and Perspectives*, *5*(2&3), 107–118.

Hill, H. C., Sleep, L., Lewis, J. M., & Ball, D. L. (2007b). Assessing teachers' mathematical knowledge. In F. K. Lester (Ed.), *Handbook for research on mathematics education* (2nd ed., pp. 111–155). Charlotte: Information Age.

Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: conceptualising and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education*, *39*(4), 372–400.

Hofstede, G. (1983). Culture's consequences: international differences in work-related values. *Administrative Science Quarterly*, *28*(4), 625–629.

Hofstede, G. (1993). Cultures and organizations: software of the mind. *Administrative Science Quarterly*, *38*(1), 132–134.

Hsieh, F.-J. (2010). Indicators of teaching ability for ideal mathematics teachers (theoretical framework). In F.-J. Hsieh & C. Chong (Eds.), *Learning from students—sailing for elaborated teaching* (pp. 9–18). Taipei: MOE.

Hsieh, F.-J., Wang, T.-Y., Hsieh, C.-J., Tang, S.-J., Chao, G.-H., Law, C.-K., et al. (2010). *A milestone of an international study in Taiwan teacher education—an international comparison of Taiwan mathematics teachers*. Taiwan TEDS-M 2008.

Hsieh, F.-J., Law, C.-K., Shy, H.-Y., Wang, T.-Y., Hsieh, C.-J., & Tang, S.-J. (2011). Mathematics teacher education quality in TEDS-M—globalizing the views of future teachers and teacher educators. *Journal of Teacher Education*, *60*(2), 172–187.

Hsieh, F.-J., Yang, C.-Z., & Shy, H.-Y. (2013). OTL of secondary mathematics teacher in teacher education curriculum. In F.-J. Hsieh (Ed.), *Taiwan TEDS-M national report*.

Huillet, D. (2009). Mathematics for teaching: an anthropological approach and its use in teacher training. *For the Learning of Mathematics: Knowing and Using Mathematics in Teaching*, *29*(3), 4–10.

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A., & Williams, C. (2008). Gender similarities characterize math performance. *Science*, *321*, 494–495.

Jakobsen, A., Thames, M. H., Ribeiro, C. M., & Delaney, S. (2012). *Using practice to define and distinguish horizon content knowledge*. Paper prepared for the twelfth International Conference on Mathematical Education, Seoul, Korea.

Johansone, I., & Malak, B. (2008). Translation and national adaptations of the TIMSS 2007 assessment and questionnaires. In J. F. Olson, M. O. Martin, & I. V. S. Mullis (Eds.), *TIMSS 2007 technical report* (pp. 63–75).

Johnson, T. P. (1998). Approaches to equivalence in cross-cultural and cross-national survey research. In J. Harkness (Ed.), *Cross cultural survey equivalence*. Mannheim: ZUMA.

Kaiser, G. (1995). Results from a comparative empirical study in England and Germany on the learning of mathematics in context. In C. Sloyer, W. Blum, & I. Huntley (Eds.), *Advances and perspectives on the teaching of mathematical modelling and applications* (pp. 83–95). Yorklyn: Water Street Mathematics.

Kane, M. T. (1992). An argument-based approach to validity. *Psychological Bulletin*, *112*, 527–535.

Kane, M. T., Crooks, T., & Cohen, A. (1999). Validating measures of performance. *Educational Measurement: Issues and Practice*, *18*(2), 5–17.

Kazima, M., & Adler, J. (2006). Mathematical knowledge for teaching: adding to the description through a study of probability in practice. *Pythagoras*, *63*, 46–59.

Kilpatrick, J., Swafford, J., & Findell, B. (Eds.) (2001). *Adding it up: helping children learn mathematics*. Washington: National Academy Press.

KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland) (Ed.) (2004). *Bildungsstandards im Fach Mathematik für den Mittleren Schulabschluss (Jahrgangsstufe 10)*. München: Wolters Kluwer.

Krauss, S., Baumert, J., & Blum, W. (2008a). Secondary mathematics teachers' pedagogical content knowledge and content knowledge: validation of the COACTIV constructs. *ZDM—The International Journal on Mathematics Education*, *40*, 873–892.

Krauss, S., Brunner, M., Kunter, M., Baumert, J., Blum, W., Neubrand, M., et al. (2008b). Pedagogical content knowledge and content knowledge of secondary mathematics teachers. *Journal of Educational Psychology*, *100*(3), 716–725.

Kristjánsdóttir, A. (2008). *Developing of teachers' professional knowledge of mathematics. Historical, present and future perspectives*. Paper presented at ICME-11, Topic Study Group 27 in Mexico.

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—The International Journal on Mathematics Education*, *44*, 371–385.

Leder, C., Pehkonen, E., & Törner, G. (Eds.) (2002). *Beliefs: a hidden variable in mathematics education?* Dordrecht: Kluwer Academic Publishers.

Leinhardt, G., & Greeno, G. (1986). The cognitive skill of teaching. *Journal of Educational Psychology*, *78*(2), 75–95.

LeTendre, G. K. (1999). The problem of Japan: qualitative studies and international educational comparisons. *Educational Researcher*, *28*(2), 38–45.

Leung, F. K. S. (2001). In search of an East Asian identity in mathematics education. *Educational Studies in Mathematics*, *47*(1), 35–51.

Leung, F. K. S., Graf, K.-D., & Lopez-Real, F. J. (2006). Mathematics education in different cultural traditions: a comparative study of East Asia and the West. In F. K. S. Leung, K.-D. Graf, & F. J. Lopez-Real (Eds.), *Mathematics education in different cultural traditions. A comparative study of East Asia and the West. The 13th ICMI study* (pp. 1–20). New York: Springer.

Ma, L. (1999). *Knowing and teaching elementary mathematics*. Mahwah: Lawrence Erlbaum Associates Inc.

Monk, D. H. (1989). The education production function: its evolving role in policy analysis. *Educational Evaluation and Policy Analysis*, *11*(1), 31–45.

Mosvold, R., Fauskanger, J., Jakobsen, A., & Melhus, K. (2009). Translating test items into Norwegian—without getting lost in translation. *Nordic Studies in Mathematics Education*, *14*(4), 101–123.

Mueller, C. W., & Parcel, T. L. (1981). Measures of socioeconomic status: alternatives and recommendations. *Child Development*, *52*, 13–30.

Mullis, I. V. S., Martin, M. O., Beaton, A. E., Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1997). *Mathematics achievement in the primary school years: IEA's third international mathematics and science study*. Chestnut Hill: Boston College.

Mullis, I. V. S., Martin, M. O., & Foy, P. (2008). *TIMSS 2007 international mathematics report: findings from IEA's trends in international mathematics and science study at the fourth and eighth grades*. Chestnut Hill: TIMSS & PIRLS International Study Center, Boston College.

Musset, P. (2010). *Initial teacher education and continuing training policies in a comparative perspective: current practices in OECD countries and a literature review on potential effects* (OECD Education Working Papers, No. 48). Paris: OECD Publishing.

Naik, S. (2008). *The measures for understanding teachers' mathematical knowledge for teaching fractions—how do they really work*? Paper presented at ICME-11, Topic Study Group 27 in Mexico.

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

National Research Council (2003). *Understanding others, educating ourselves: getting more from international comparative studies in education*. Washington: National Academies Press.

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.

Niss, M. (2002). *Mathematical competencies and the learning of mathematics: the Danish KOM project*. IMFUFA, Denmark: Roskilde University.

Paine, L. (1997). Chinese teachers as mirrors of reform possibilities. In W. K. Cummings & P. G. Altbach (Eds.), *The challenge of eastern Asian education* (pp. 65–83). Albany: SUNY Press.

Paine, L., & Ma, L. (1993). Teachers working together: a dialogue on organizational and cultural perspectives of Chinese teachers. *International Journal of Educational Research*, *19*(8), 675–697.

Pelgrum, W. J., Eggen, T., & Plomp, T. (1986). *Second international mathematics study: the implemented and attained mathematics curriculum—a comparison of eighteen countries*. Washington: Center for Education Statistics.

Pepin, B. (1999). Existing models of knowledge in teaching: developing an understanding of the Anglo/American, the French and the German scene. In B. Hudson, F. Buchberger, P. Kansanen, & H. Seel (Eds.), *TNTEE publications: Vol. 2(1). Didaktik/Fachdidaktik as science(s) of the teaching profession* (pp. 49–66).

Perry, N. E., Phillips, L., & Hutchinson, L. R. (2006). Preparing student teachers to support for self-regulated learning. *Elementary School Journal*, *106*, 237–254.

Reckase, M. D. (2009). *Multidimensional item response theory*. New York: Springer.

Ribeiro, A. J. (2008). *The multimeanings of equation: considerations concerning its potentialities in teaching and learning of algebra*. Paper presented at the 11th International Congress on Mathematical Education, Monterrey, Mexico.

Richardson, V. (1996). The role of attitudes and beliefs in learning to teach. In J. Sikula, T. Buttery, & E. Guyton (Eds.), *Handbook of research on teacher education* (2nd ed., pp. 102–119). New York: Macmillan.

Rowland, T. & Ruthven, K. (Eds.) (2010). *Mathematics education library: Vol. 50. Mathematical knowledge in teaching*. Berlin: Springer.

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*, 255–281.

Sadler, D. R. (2012). Making competent judgements of competence. In S. Blömeke, O. Zlatkin-Troitschanskaia, Ch. Kuhn, & J. Fege (Eds.), *Modeling and measuring competencies in higher education. Tasks and challenges*. Rotterdam: Sense Publishers.

Scheerens & Bosker (1997).

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), 118–127.

Schmidt, W. H., McKnight, C. C., Valverde, G. A., Houang, R. T., & Wiley, D. E. (1997). *Many visions, many aims: a cross-national investigation of curricular intentions in school mathematics*. Dordrecht: Kluwer.

Schmidt, W., Houang, R., Cogan, L., Blömeke, S., Tatto, M. T., Hsieh, et al. (2008). Opportunity to learn in the preparation of mathematics teachers: its structure and how it varies across six countries. *ZDM—The International Journal on Mathematics Education*, *40*(5), 735–747.

Schmidt, W. H., Blömeke, S., & Tatto, M. T. (2011). *Teacher education matters. A study of the mathematics teacher preparation from six countries*. New York: Teacher College Press.

Schön, D. (1983). *The reflective practitioner, how professionals think in action*. New York: Basic Books.

Schoenfeld, A. H. (1994). Reflections on doing and teaching mathematics. In *Stud. math. think. learn. Mathematical thinking and problem solving* (pp. 53–75). Hillsdale: Lawrence Erlbaum.

Schoenfeld, A. H. (2010). *How we think: a theory of goal-oriented decision making and its educational applications*. New York: Routledge.

Schoenfeld, A. H., & Kilpatrick, J. (2008). Toward a theory of proficiency in teaching mathematics. In D. Tirosh & T. Wood (Eds.), *International handbook of mathematics teacher education, volume 2: tools and processes in mathematics teacher education* (pp. 321–354). Rotterdam: Sense Publishers.

Shavelson, R. (2012). An approach to testing & modeling competence. In S. Blömeke, O. Zlatkin-Troitschanskaia, Ch. Kuhn, & J. Fege (Eds.), *Modeling and measuring competencies in higher education. Tasks and challenges*. Rotterdam: Sense Publishers.

Shulman, L. (1985). Paradigms and research programs in the study of teaching: a contemporary perspective. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 3–36). New York: Macmillan.

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 Research*, *57*, 1–22.

Simmons, P. (1995). Gender, ethnic and socio-ethnic differences in attainment and progress: a longitudinal analysis of student achievement over 9 years. *British Educational Research Journal*, *21*, 465–483.

Singh, K., Granville, M., & Dika, S. (2002). Mathematics and science achievement: effects of motivation, interest and academic engagement. *Journal of Educational Research*, *95*, 323–332.

Stark, J., & Lattuca, L. R. (1997). *Shaping the college curriculum: academic plans in action*. Boston: Allyn and Bacon.

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: best ideas from the world's teachers for improving education in the classroom*. New York: The Free Press.

Tatto, M. T., Schwille, J., Senk, S., Ingvarson, L., Peck, R., & Rowley, G. (2008). *Teacher education and development study in mathematics (TEDS-M): policy, practice, and readiness to teach primary and secondary mathematics. Conceptual framework*. East Lansing: Teacher Education and Development International Study Center, College of Education, Michigan State University.

Tatto, M. T., Lerman, S., & Novotna, J. (2010). *The organization of the mathematics preparation and development of teachers: a report from the ICMI Study 15*. Published online at Springer Science and Business Media B.V.

Tatto, M. T., Schwille, J., Senk, S. L., Ingvarson, L., Rowley, G., Peck, R., et al. (2012). *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries: findings from the IEA teacher education and development study in mathematics (TEDS-m)*. Amsterdam: IEA.

Thomas, W. P., & Collier, V. P. (1997). *School effectiveness for language minority students*. National Clearinghouse for English Language Acquisition.

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

Travers, K. J. & Westbury, I. (Eds.) (1989). *The IEA study of mathematics I: analysis of mathematics curricula*. Oxford: Pergamon.

UNESCO (1997). *International standard classification of education—ISCED*. Paris: UNESCO.

Vandenberg, R. J., & Lance, Ch. E. (2000). A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*, 4–70.

Van de Vijver, F. J. R. (1998). Towards a theory of bias and equivalence. In J. A. Harkness (Ed.), *Zuma-Nachrichten, Spezial: Vol. 3*. *Cross-cultural survey equivalence* (pp. 41–66). Mannheim: ZUMA.

van Ewijk, R., & Sleegers, P. (2010). The effect of peer socioeconomic status on student achievement: a meta-analysis. *Educational Research Review*, *5*, 134–150.

Walter, O., & Taskinen, P. (2008). Der Bildungserfolg von Jugendlichen mit Migrationshintergrund in den deutschen Ländern. In M. Prenzel, C. Artelt, J. Baumert, W. Blum, M. Hammann, E. Klieme, & R. Pekrun (Eds.), *PISA 2006 in Deutschland. Die Kompetenzen der Jugendlichen im dritten Ländervergleich* (pp. 343–374). Münster: Waxmann.

Wang, J., & Lin, E. (2005). Comparative studies on US and Chinese mathematics learning and the implications for standards-based mathematics teaching reform. *Educational Researcher*, *34*(5), 3–13.

Weinert, F. E. (2001). Concepts of competence: a conceptual clarification. In D. S. Rychen & L. H. Salgnik (Eds.), *Defining and selecting key competencies* (pp. 45–66). Göttingen: Hogrefe.

West, K. I. (2008). *Japanese high school mathematics teacher competence in real world problem solving*. Paper presented at ICME-11, Topic Study Group 27 in Mexico.

Wigfield, A., & Eccles, J. S. (2000). Expectancyvalue theory of achievement motivation. *Contemporary Educational Psychology*, *25*, 68–81.

Wilson, S. M., Floden, R. E., & Ferrini-Mundy, J. (2001). *Teacher preparation research. Current knowledge, gaps, and recommendations*. Washington: Center for the Study of Teaching and Policy.

Wu, H. H. (1999). On the education of mathematics majors. In *MSRI: Vol. 36*. *Contempory issues in mathematics education*.
Wu, Z., & Li, L. (2008). *Sharpening teaching ability in K-8 mathematics classrooms*. Paper presented at ICME-11, Topic Study Group 24 in Mexico.
Wu, Z. (2010). *Mathematics teacher readiness in multiculturally and linguistically diverse classrooms: a new approach of California teacher performance assessment*. Unpublished Presentation.
Zeichner, K. (1980). Myth and realities: field based experiences in pre-service teacher education. *Journal of Teacher Education*, *31*(6), 45–55.