Int J of Sci and Math Educ (2015) 13:331–350
DOI 10.1007/s10763-015-9618-5

Early Career Mathematics Teachers' General Pedagogical Knowledge and Skills: Do Teacher Education, Teaching Experience, and Working Conditions Make a Difference?

Johannes König · Sigrid Blömeke · Gabriele Kaiser

Received: 3 May 2014 / Accepted: 31 December 2014 / Published online: 20 March 2015

Abstract We examined several facets of general pedagogical knowledge and skills of early career mathematics teachers, asking how they are associated with characteristics of teacher education, teaching experience, and working conditions. Declarative general pedagogical knowledge (GPK) was assessed via a paper-and-pencil test, while early career teachers' skills to perceive and interpret classroom situations were assessed via video-vignettes. Data from a follow-up study of TEDS-M Germany in 2012 were used, including a sample of 278 early career middle school teachers of mathematics. While teachers' declarative knowledge can be predicted by teacher education grades, teachers' skill to interpret classroom situations presented by videos can be predicted by their amount of time spent on teaching relative to their overall working time, which is interpreted as a form of deliberate practice. Different competence profiles of pedagogical knowledge and skills are identified via latent-class analysis. Besides teaching experience, profiles are associated with generic teaching challenges (motivating students, disruptive student behaviour) perceived by the teachers. Implications of findings for professional development of early career teachers are discussed.

Keywords Assessment · Competence · General pedagogical knowledge · Teacher · Teacher education · Video-vignettes · Latent-class analysis · Competence profile

Currently, research on the development of teacher competence is of great interest. Standardised instruments have been developed to assess the knowledge and skills of

Electronic supplementary material The online version of this article (doi:10.1007/s10763-015-9618-5) contains supplementary material, which is available to authorized users.

J. König (✉)
University of Cologne, Cologne, Germany
e-mail: johannes.koenig@uni-koeln.de

S. Blömeke
University of Oslo, Oslo, Norway

G. Kaiser
University of Hamburg, Hamburg, Germany

pre- and in-service teachers (cf. Baumert, Kunter, Blum, Brunner, Voss, Jordan & Tsai, 2010; Hill 2010; Kersting, Givvin, Thompson, Santagata & Stigler, 2012; Tatto, Schwille, Senk, Ingvarson, Peck & Rowley, 2008). Researchers draw on a shared model of teacher competence, according to which the different knowledge and skill facets are related to the requirements of successful teaching (Blömeke, Gustafsson & Shavelson, 2015; Kaiser, Benthien, Döhrmann, König & Blömeke, 2013; Schoenfeld, 2011; Shavelson, 2010; Shulman, 1987; Weinert, 2001). Teacher competence is regarded as a multidimensional construct, consisting of content knowledge (CK), pedagogical content knowledge (PCK) and general pedagogical knowledge (GPK) as well as of perception, interpretation and decision-making skills. Especially the research on teacher expertise has worked out that both knowledge and skills contribute to the expert's performance in the classroom (Bromme, 2001).

Researchers have started to develop instruments that assess these different modes. In 2012, a follow-up study (TEDS-FU) of the *Teacher Education and Development Study—Learning to Teach Mathematics* (TEDS-M; Tatto et al., 2008) was carried out in Germany, sampling TEDS-M participants now in the stage of early career teachers (i.e. 4 years teaching experience or less).[1] The TEDS-M tests for examining mathematics teachers' knowledge in the area of CK, PCK and GPK were applied again, but they were extended by video-vignettes assessing the skills to perceive and interpret generic pedagogical and pedagogical-content-related challenges occurring in real classroom situations. Thus, the knowledge and skill facets of early career mathematics middle school teachers' competence can be examined. This article will have a special focus on mathematics teachers' GPK as assessed by the (digitalised) paper-pencil test and their general pedagogical skills to perceive and interpret as assessed by the video-vignette test.

Early career teaching—the major challenge of the teachers examined by TEDS-FU—has been considered a particularly crucial phase of teachers' professional development (cf., e.g. McCormack, Gore & Thomas, 2006). Beginning teachers are challenged with the various, rather new experiences they gain through everyday teaching, especially when working full time (Keller-Schneider & Hericks, 2011). While, in some countries such as Taiwan, the teacher education system explicitly supports beginning teachers with an induction phase followed by continuous professional development activities, other countries such as Germany seem to offer less support. After finishing initial teacher education, which consists of a first phase at university with heavy focus on theory and a second practical phase that can be compared with other countries' induction programs, beginning teachers in Germany are no longer systematically required to engage in further professional development courses.

However, research on the development of teacher expertise shows the way to become an expert teacher is much longer than just passing initial teacher education examinations. Instead, several years of in-service teaching and deliberate practice are essential (Berliner, 2004; Ericsson, Krampe & Tesch-Romer, 1993). Furthermore, the development of expertise is very complex and varies across individuals. The knowledge acquired during teacher education is consolidated and transformed in individually different ways when teachers start to work in different school contexts (McCormack et al., 2006). The accumulation of experience from which teachers develop mental

[1] TEDS-FU was funded by the German Research Foundation (DFG, BL 548/8-1). The views expressed in this paper are those of the authors and do not necessarily reflect those of the DFG.

models and strategies (Dehoney, 1995) varies between individuals to a large extent. Early career teachers are confronted with a variety of challenges, and individual dispositions, teaching experience, student composition to deal with, and perceived quality of working conditions of the individual school may vary. Thus, the way to become a professional teacher can differ considerably for different individuals, even within broadly similar cultural contexts (e.g. country) and settings (e.g. school type and subject).

Against this background, in this article, we examine the question how the GPK of early career (i.e. 4 years teaching experience or less) mathematics middle school teachers and their skills to perceive and interpret pedagogical situations in a mathematics classroom presented to them via video-vignettes are influenced by variables related to teacher education, teaching experience, and teachers' working conditions. Thus, we look at an important career step of teachers in which their expertise is under development. We ask how several general pedagogical facets of the teachers are related to what they experience and are confronted with amidst their starting years as in-service teachers. We assume that it is very important how those factors may have an influence on the individual teacher's acquisition and further development of general pedagogical knowledge and skills.

Conceptual Framework

Defining General Pedagogical Knowledge

In TEDS-M 2008, a theoretical framework of teachers' GPK that could be transformed into a paper–pencil instrument and be tested empirically across countries was developed (see, for details, König, Blömeke, Paine, Schmidt & Hsieh, 2011). Following the notion of "competence" (Shavelson, 2010; Weinert, 2001), the framework focussed on the mastering of professional tasks and its underlying latent cognitive dispositions. Instruction was identified as the core activity of teachers in all subjects and countries (Berliner, 2001, 2004; Bromme, 1992). The international state of instructional research served therefore as the rationale to select topics and cognitive demands to be covered in the GPK test.

Instructional models used across countries to describe effective teaching (see, for details, König et al., 2011) provided four generic dimensions of teaching responsibilities: to prepare, structure and evaluate lessons ("structure"); to motivate and support student learning as well as to manage the classroom ("motivation/classroom management"); to deal with heterogeneous learning groups in the classroom ("adaptivity"); and to assess student achievement ("assessment"). In addition, three dimensions of cognitive processes describing the cognitive demands on teachers when dealing with such generic classroom situations were defined following Anderson & Krathwohl (2001): to retrieve information from long-term memory in order to describe the classroom situation; to understand or analyse a concept, a specific term or a phenomenon outlined; and to generate strategies for how they would solve the problem posed (for more details, see König et al., 2011). Generic teaching responsibilities and cognitive demands made up a matrix which served as a heuristic for the development of items intended to assess GPK.

Perceiving and Interpreting Classroom Situations

The challenge to measure context-dependent skills of teachers makes it necessary to go beyond the limited scope of paper–pencil assessments (Shavelson, 2010). Therefore, these situated facets of teacher competence were assessed via video-vignettes in the follow-up study of TEDS-M carried out in Germany (TEDS-FU; König, Blömeke, Klein, Suhl, Busse & Kaiser, 2014). Irrespective of the field or the country context, the skills to notice, that means to perceive and interpret classroom situations, are considered to be an important premise for the successful mastering of teaching (van Es & Sherin, 2002; Sherin, Jacobs & Philipp, 2011; Blömeke et al., 2015), leading to the basic distinction between perceiving and interpreting.

Perceiving

From the research on teacher expertise, which has proven to be valid across different subjects and countries, it is well known that expert teachers outperform novice teachers in recalling meaningful instructional details (Klein & Hoffman, 1993; König & Lebens, 2012). Expert teachers' categorical perception with which phenomena, events or sequences are cognitively divided into relevant units for perception (e.g. Bromme, 1992) supports them to focus on the *relation* between knowledge elements rather than on discrete elements. Repeated activation of schemata strengthens connections between elements within a schema and support enhanced activation of knowledge for categorising new information when salient cues are present. Since connectivity and complexity of schemata required for identifying and categorising information evolve with practice (e.g. Dehoney, 1995), perceptual accuracy is an indicator of expertise. Consequently, it can be reasonably assumed that expert teachers identify relevant instructional situations seen in a video-vignette assessment more precisely and correctly than do novices (Sabers, Cushing & Berliner, 1991).

Additionally, expert teachers can be characterised by a more holistic perception compared to novices (Bromme 2001; König & Lebens, 2012): They reconstruct and anticipate the context of instruction and engage in reflecting alternative problem-solving strategies. Whereas novice teachers observe classroom situations step by step due to the fragmented structure of their knowledge, experts have an intuitive grasp of the situation since their knowledge is highly interlinked (Bromme, 1992). More specifically, prior knowledge of experts organised in schemata is employed during perception to form a cognitive representation of the situation (Putnam, 1987). By contrast, novices, whose knowledge structures for constructing a mental framework have not yet been developed, are likely to experience difficulties in reconstructing the context of instruction.

Interpreting

The functional interpretation of instructional events and sequences depends on reasoning about the instructional intention and rationale amidst the context of classroom teacher-student interaction. Although this functional interpretation of actions is rarely explicated in everyday teaching situations, it can be accessed from long-term memory (Bromme, 1992). In contrast to teachers' holistic perception, the

interpretation of events goes beyond generating mental representations, since it strongly depends on reframing and transforming knowledge. Whereas the holistic perception can be described as a perceptive-representational process, the interpretation of events refers to transformative processes.

The ability to transform knowledge is essential to derive an interpretation and to verbalise it explicitly (cf. Hackl, 2004). When expert teachers are provided with video-vignettes, they are more likely to come up with functional interpretations than novices (Berliner, 1992). The depth of information processing serves as an explanatory model for these differences, since experts draw conclusions about meaning and reasons behind instructional sequences (e.g. "students work in groups because solving this specific task requires cooperation"). Thus, the quantitative amount and the qualitative structure of prior knowledge gives an advantage to expert teachers who are able to quickly interpret a given classroom situation as a whole and to draw conclusions even about the non-visible features such as the background and instructional rationale behind the perceived scenario.

Early Career Teachers' Opportunities to Learn

The state of research does not tell us precisely how opportunities to learn (OTL) in different kinds of teacher education programs or professional development initiatives are related to facets of early career teachers' general pedagogical knowledge and skills. This is mainly due to the lack of studies assessing these in a standardised way (König & Blömeke, 2012; König et al., 2014). Many teacher education programs worldwide intend future teachers acquire professional pedagogical knowledge and skills (Tatto et al., 2008). Thus, general pedagogical OTL are provided by teacher education institutions and initiatives of professional development in many countries. While courses in the academic setting often primarily aim at the acquisition of theoretical knowledge, in-school OTL give future teachers the chance to connect their knowledge to practical situations in the classroom (König & Blömeke, 2012; König et al., 2014).

Regarding GPK as defined for the TEDS-M assessment, such practical experiences whereby pre-service teachers or early career teachers have the chance to teach students in the classroom should be important. Pre-service as well as in-service teachers are forced to reflect on tasks such as structuring lessons, dealing with heterogeneity, or motivating students prior, during and after the teaching process, and thus to activate their GPK previously acquired in the academic context of teacher education. Presumably, while making use of GPK in such situations, future as well as early career teachers become also increasingly flexible in how to apply their knowledge to different classroom contexts, starting to develop their skills to perceive and interpret classroom situations (Berliner, 2001, 2004; Gruber & Rehrl, 2005). Thus, we conclude that teaching experience may be a significant factor for the development and further acquisition of early career teachers' general pedagogical knowledge and skills, especially when linked with teachers' deliberate efforts to reflect on and improve their teaching (Ericsson et al., 1993).

Perceived quality of early career teachers' working conditions should be another decisive factor. Occupational and teacher-related research indicates appraisal is an important feature of support that affects work quality (Kouzes & Posner, 1999). The role of the principal seems to be particularly important in this context (Valentine, Clark, Hackmann & Petzko, 2004). Self-reported data of a sub-sample of early career middle school mathematics teachers included in this study on how they regarded themselves able to cope with

the challenges of instruction and how satisfied they were with their job was significantly positively related to the level of appraisal they reported (Blömeke & Klein, 2013).

Multi-dimensional Profiles of Teacher Competence

Blömeke et al. (2015) suggest viewing competence as a continuum that includes different types of resources including not only more trait-like dispositions such as CK, PCK or GPK but also more performance-oriented skills such as perceiving, interpreting or decision making—all relevant in the transformation of resources into performance. They hypothesise that different persons may dispose of qualitatively different *profiles* of these resources. One important research question in this context is therefore how the different resources are linked to each other, what this interplay depends on and how the resources can be built up.

Since research on this question does not yet exist, it is difficult to hypothesise the precise shape of these profiles. Latent-class analysis (LCA) provides a tool to explore such unobserved heterogeneity in a group. LCA represents a relatively new but increasingly popular method in item-response theory that is person-oriented in contrast to the traditional variable orientation. It is a model-based method and was specifically developed to identify distinct profiles. The classes identified represent subpopulations where population membership was inferred from the data (Magidson & Vermunt, 2004). The person-oriented perspective is especially valuable in the complex field of competence assessment because it is able to integrate several facets and their interactions, thus providing a more holistic picture (von Eye & Bergman, 2003).

Research Questions and Hypotheses

In this article, we examine facets of general pedagogical knowledge and skills of early career mathematics teachers, asking how they are associated with variables of teacher education, teaching experience, and working conditions. These three components are chosen, since they together constitute most of the setting which should be accounted for when describing and analysing how early career teachers acquire and develop knowledge and skills related to general pedagogy. They are also regarded to be important because they basically depend on the teacher education and school system, thus being decisive factors for supporting the professional development of early career teachers.

Data from TEDS-FU carried out in 2012 will be analysed with respect to two main research questions. First, how are the current general pedagogical knowledge and skills of early career mathematics teachers associated with variables of teacher education, teaching experience, and working conditions? Second, can general pedagogical competence profiles be identified and related to these variables?

The first question takes on the traditional variable-oriented approach and is related to differences in teacher education, teaching experience and working conditions, which may influence the professional development of early career teachers reflected in their current level of pedagogical knowledge and skills. Findings from TEDS-M had revealed that the more advanced future teachers were in the course of their initial teacher education, the better they performed on the test measuring GPK (König, 2013).

Continuing the examination in this field, additional analysis showed that practical in-school experience served as a relevant opportunity to learn, too, and this with respect to

future primary teachers in Germany and the US (König & Blömeke, 2012): Future teachers who had an appropriate balance between teaching and being supported by mentors and who had had opportunity to reflect on and improve their teaching to a relatively large extent outperformed other future teachers who either had been less supported by mentors or who had gained only little teaching experience. Since initial teacher education explicitly aims at fostering future teachers' ability to reflect on teaching and their professional knowledge in the area of general pedagogy, one can conclude from these findings that the knowledge tested is curricular valid with regards to initial teacher education. Teacher education *grades* indicating the achievement at the end of the first and second phases of German teacher training, therefore, should significantly predict the scores that early career teachers reach in the GPK test (we denote this as hypothesis 1a, in the following abbreviated as H1a).

Moreover teaching type, i.e. whether they have been qualified to teach in middle school only or, additionally, to teach in upper secondary school, should have a significant influence on pedagogical performance (H1b): *Teaching type* indicates whether an early career teacher has the qualification to teach mathematics in grades 1 or 5 through 10 (elementary and middle school or middle school only) or in grades 5 through 12 (middle school and high school). In Germany, this distinction corresponds with the stratification of the school system into different types of middle school (*Hauptschule* and *Realschule* or their combinations), on the one hand, and grammar school (*Gymnasium*), on the other hand (see, for details, König & Blömeke, 2013). Since at grammar schools, a heavy focus is on subject-related learning, usually early career teachers qualified for the *Gymnasium* have had a large amount of opportunities to learn related to their subjects and sometimes slightly less opportunities to learn in general pedagogy compared with those qualified for middle schools specifically. Middle school teacher education has, in turn, a special focus on pedagogical knowledge and skills. Therefore, our variable "teaching type" reflects important priorities related to initial teacher education program and career paths characteristics.

Regarding the pedagogical video-vignettes assessing pedagogical skills, teachers' test scores presumably will depend on their teaching experience and the perceived quality of their working conditions rather than on initial teacher education. We assume that early career teachers who have already become increasingly flexible in how to apply their knowledge and who have increasingly developed perceptual and interpretation skills while repeatedly and extensively mastering the challenges during teaching will have higher test scores. Teaching experience, especially the intensive and continuous reflection on one's own teaching practice, supports early career teachers' competence development (Schön, 1983; Hart, Alston & Murata, 2011).

Regarding the component of teaching experience, in our analysis, three hypotheses will be examined. First, we hypothesise that early career teachers working full time will do better in the test than those working part time (H1c), since they *generally* gain more experiences through everyday teaching compared with teachers working part time only (Keller-Schneider & Hericks, 2011). Second, we hypothesise that the amount of time spent on preparation, conducting and analysis of teaching *in relation to the overall time* spent on teachers' everyday work is positively associated with the improvement of pedagogical skills (H1d) because such a prioritising of time may at least partly reflect a teacher's deliberate efforts to improve his or her teaching. Since pedagogical skills are not limited to one subject only, we assume that extensive and varying teaching

experiences, i.e. the experience of teaching additional subjects others than mathematics, will also contribute to early career teachers' competence development (H1e), since becoming an expert teacher depends on the continuous working through the complexities of teaching, which is not reduced to problems or situations of one subject only that may be handled with routine procedures (cf. Dunn & Shriner, 1999).

The professional development of teachers presumably depends on the quality of their working conditions. In our study, we focus on two aspects assuming they are relevant when explaining teachers' test scores, namely demand and reward of teaching tasks. Reward in the form of appraisal should turn out to support a teacher's general pedagogical knowledge and skills (see section "Early Career Teachers' Opportunities to Learn"), letting us assume a positive correlation (the more appraisal, the higher the test scores; H1f).

By contrast, it is difficult to determine how high or low demand of teaching tasks is correlated with pedagogical knowledge and skills. On the one hand, those teachers could be overwhelmed when perceiving a too high demand of teaching tasks whose competence is limited (cf. Hobfoll & Freedy, 1993), i.e. who show lower test scores. On the other hand, those teachers who perceive a high demand of teaching tasks may need pedagogical knowledge and skills to be able to perceive and interpret the classroom situations appropriately so that teachers with higher test scores also report higher demands. Furthermore, these cannot be mastered by just applying routines but may stimulate deliberate practice, which finally could result in higher test scores. Due to concurrent assumptions about the direction of correlations between demand of teaching tasks and test scores, in this case, we waive to define a concrete hypothesis.

Our second research question takes on the rather new person-oriented approach and is related to the question whether competence profiles of early career middle school teachers can be identified using general pedagogical knowledge and skills as indicators and whether these profiles then can be explained by teacher education, teaching experience, and working conditions indicators. We hypothesise that such profiles can be identified among early career teachers (H2a). In analogy to the hypotheses of our first research question, we assume that these profiles are associated with teacher education grades, teaching experience, and working conditions (H2b).

To control for background variables of teachers, in our analyses, we include sex and age.

Method

Sample

In Germany, the TEDS-M follow-up study was carried out in 2012 (TEDS-FU). TEDS-FU sampled German early career middle school teachers of mathematics who had participated in TEDS-M and who had agreed to be followed up a few years later. The TEDS-M 2008 target population on which therefore TEDS-FU is based was defined as "future teachers who are in their final year of training before they are eligible to become practicing teachers of mathematics (…) in lower secondary schools (either as generalist teachers or as mathematics specialists)" (Tatto et al., 2008, p. 32). In Germany, the sample of middle school mathematics teachers consisted of future teachers attending teacher education programs that qualified them either as elementary and middle school

or pure middle school mathematics teachers (allowing them to teach from grade 1 or 5 through 10) or as middle school and high school mathematics teachers (allowing them to teach from grade 5 through 12). Seven hundred seventy-one future middle school teachers from Germany were surveyed in TEDS-M out of which about 400 had agreed to be followed up.

In TEDS-FU, all data collection was carried out online due to difficulties in reaching the former TEDS-M participants now working at schools spread across Germany. In TEDS-FU, first, an online survey was conducted including 278 middle school mathematics teachers that formerly had participated in TEDS-M (36.1 % response rate in relation to the representative TEDS-M sample, about 70 % response rate in relation to those who had agreed to be followed up).[2] That component consisted of a questionnaire covering teachers' background information, vocational status, teaching experience and workplace conditions. As another survey component, teachers were asked to take an online assessment conducted via a web-browser interface with password requirement (not via email) and blocks of items were timed, thus preventing participants from cheating. This component comprised the (digitalised) paper–pencil test developed in TEDS-M to measure GPK (König et al., 2011) and video-vignettes providing typical classroom situations as a stimulus followed by various test items to measure pedagogical skills to perceive and interpret classroom situations (König et al., 2014). In total, 171 teachers participated in this second component (22.2 % TEDS-M response rate; about 45 % response rate in relation to those who had agreed to be followed up).

On average, the teachers in our sample were about 32 years old (M=31.9, SD=5.5). Almost 60 % were female. About 47 % were teachers for grades 1 or 5 through 10, whereas about 53 % were teachers for grades 5 through 12. Although these characteristics fit well to the TEDS-M characteristics, the TEDS-FU sample has to be regarded as a convenience sample. Due to a two-step self-selection process (first, agreement to being followed up had to be signed; second, the teachers had then in fact to be willing to participate in TEDS-FU after being followed up), the sample is biased towards teachers with stronger GPK (see, for details and a discussion for limitations of the sample, König et al., 2014).

Assessment Instruments

Paper-Pencil Test Measuring General Pedagogical Knowledge

As laid out above, generic dimensions of teaching responsibilities and cognitive demands made up a matrix, which served as a heuristic for the development of GPK items in TEDS-M. For each cell, a subset of items was developed (for details of test development, see König et al., 2011). Two item examples (see the Supplementary material of this article) may illustrate the GPK test and the heuristic used to conceptualise GPK. The first item measured knowledge about "motivating" students. Future teachers had to "recall" basic terminology of achievement motivation ("intrinsic motivation" and "extrinsic motivation"), and they were asked to analyse five statements against the background of this distinction. Statement C represented an example of

[2] In contrast to Blömeke & Klein (2013), we could include additional teachers because they had data on at least one of the measurements used in this paper.

"intrinsic motivation", whereas A, B, D and E were examples for "extrinsic motivation". The second item example was an open-response item. Here, future teachers were asked to support another future teacher and evaluate her lesson. This is a typical challenge during a peer-led teacher education practicum, but practicing teachers are also regularly required to analyse and reflect on their own as well as their colleagues' lessons. The item measured knowledge of "structuring" lessons. The predominant cognitive process was to "generate" fruitful questions.

Video-Vignette Test Measuring General Pedagogical Skills

In TEDS-FU, three 3- to 4-min video-vignettes were developed and provided to early career teachers of mathematics (see, for details, König et al., 2014). The clips were developed using scripts that covered critical incidents and the full range of typical teaching steps during a lesson. Their length was reduced to 3 to 4 min due to time constraints during data collection while still being able to cover the scripted intentions. After each video clip, teachers had to respond to several test items. Items measuring perception were multiple-choice items with a Likert-rating scale (four categories) ranging from "completely agree" to "completely disagree". They were developed based on Clausen, Reusser & Klieme (2003). About half of them were related to the *precision* of teacher perception (e.g. "The teacher presents the lesson's task visually AND acoustically"); the other half measured teacher perception *holistically* (e.g. "Most students take an active part in the lesson"). The teachers had to mark the extent of their agreement. Each statement referred to the teaching in one video clip specifically. An expert rating with 26 academic and practical experts from universities and teacher training seminars was carried out to define which category of each item's rating scale was to be scored as "correct", i.e. the experts confirmed the content validity of the statements and decided which rating could be accepted as correct (final agreement, 85 %). These answers were used as benchmarks for classifying the middle school teachers' responses (see, for details, Kaiser et al., 2013).

Items measuring the skill to interpret general pedagogical classroom situations were open-response items that would allow teachers to provide cognitively more complex statements. The item in the Supplementary material is related to one of the video clips resulting from a mathematics lesson on computing the volume of a box. During that lesson, students had to work in pairs, whereby three pairs and their way of cooperation were focussed on more closely in the video clip. The item example now asks the teachers for an in-depth analysis of the three student pairs' cooperation from a pedagogical perspective.

Teacher Education, Teaching Experience, and Working Conditions Variables

To account for teachers' qualification derived from initial teacher education, we use two indicators: teacher education grades and teaching type. *Teaching type* indicates whether the single early career teacher has the qualification to teach mathematics in grades 1 or 5 through 10 and coded as 0 or in grades 5 through 12 at grammar schools (*Gymnasium*) coded as 1. *Teacher education grades* were measured by self-reports. In Germany, two grades are decisive and therefore captured: The average grade of the first state examination after completion of university and the average grade of the second state

examination after completion of the practical training phase. They usually range from 1.0 (very good) to 4.0 (barely passed).

Teaching experience is captured by teaching time, teaching experience related to subjects other than mathematics, and teaching status. To operationalise *teaching time*, teachers were asked to indicate the proportion of time spent on teaching and the preparation of teaching in relation to other tasks such as additional curricular activities. Since, in Germany, teacher tasks are not only limited to teaching but are also related to overarching tasks such as contributing to school improvement, this indicator reflects whether an early career teacher had the chance to pay attention to teaching primarily or whether he or she had to fulfill extra requirements of his or her school that may have hindered the teacher to master specific early career teaching challenges.

As our target group are mathematics teachers, their *experience to teach other subjects than mathematics* may vary. However, since, in this analysis, we test knowledge and skills that are irrespective of the single subject, such teaching experience may influence scores in the tests applied. To operationalise this aspect of teaching experience, teachers were asked to indicate whether they have taught other subjects than mathematics in grades 5 and 6, 7 and 8, and 9 and 10. Response format was dichotomous (yes=1, no=0). The three items were summed up to have one reliable indicator (α=.76). The *teaching status* of the teachers was operationalised by one dichotomous variable only, indicating whether an early career teacher had been working full time in the school years 2009/2010, 2010/2011 and 2011/2012 (coded as 1) or not (coded as 0). Of our sample, 65.3 % had been working full time over the period of the three school years.

Two scales are used to consider demand and reward of teaching tasks to describe early career teachers' working conditions. To capture teaching task demand, teachers were asked to report on *challenges* they perceived such as managing student behaviour or motivating students. A scale comprising three items was constructed. The introductory question was as follows: "What are some of the difficulties or challenges that you have encountered in your current teaching position?" Teachers had to respond to challenges (such as "manage a classroom", "motivate students") using Likert scales ranging from a "major problem" to "not a problem". The three items make up a reliable scale (α=.69). The higher a teacher rates on that scale, the less he or she reports being challenged.

Task reward was operationalised by *appraisal* teachers received from the school principal and/or the teachers' colleagues. "How often have you received appraisal and/or feedback from the following people about your work as a teacher?" was the introductory question, and for the scale in our analysis, we used two items, namely "school administration" and "other teachers", which had to be responded to on a six-point Likert scale from "never" to "more than once a month". The two items make up a reliable scale (α=.67). The higher a teacher rates on that scale, the more he or she reports having received appraisal.

Data Analyses

Item-response theory (IRT)-based scaling was applied using the software *ConQuest* (Wu, Adams & Wilson, 1997), since it provides valuable insight into the properties of the test instruments, which could not be delivered by applying methods based on classical test theory (CTT). A multi-dimensional IRT model specifying three latent

abilities (GPK as assessed via the digitalised paper-and-pencil test, the skill to perceive and the skill to interpret classroom situations as assessed via video-vignettes) turned out to fit the data better than a model in which only one latent variable was specified by all test items (see, for details, König et al., 2014). Thus, in the following, we use three indicators to describe early career mathematics teachers' competence: (1) general pedagogical knowledge and skills to (2) perceive and (3) interpret classroom situations.

In the following analysis, data from the two study components (online survey of teacher education characteristics and context conditions as well as the online assessment of knowledge and skills, see "Sample" for more details) are used. All data analysis will be carried out using the Software *Mplus* (Muthén & Muthén, 1998–2006), which enables us to adequately deal with missing data through model-based imputation of data (*full-information-maximum-likelihood* option).

To answer our first research question and as a first step of data analysis, multiple regression analysis was carried out in which the three facets of early career teachers' pedagogical competence were specified as dependent and several predictors related to the teachers' background (as control variables), teacher education, teaching experience and working conditions were specified as independent variables (see Table 1). All regression coefficients were estimated simultaneously; thus, effects were controlled for inter-correlations between criteria variables.

To answer our second research question, LCA was applied to identify classes which group together persons who share similar characteristics or behaviour (Magidson & Vermunt, 2004). In our case, we are interested to model pedagogical profiles among teachers and generate profiles using knowledge and the skills to perceive and interpret as manifest indicator variables. Latent-class models for continuous data were computed using the software *Mplus* (Muthén & Muthén, 1998–2006). The decision about the number of classes was based, firstly, on the information criteria Akaike's Information Criterion (AIC) and the Bayesian Information Criterion (BIC) that allow the decision on which model is to be preferred (the smaller the estimate, the better the model fits the data). Secondly, we evaluated the classification quality based upon entropy as an aggregated uncertainty measure. Finally, the estimate of the mean probability for a teacher's most likely latent class membership and the step parameters were examined. This showed that the item steps were sorted as expected which is an important quality criterion (Rost & Georg, 1991).

Results

Variable-Centred Analysis: Regression of General Pedagogical Facets on Predictors

In general, the regressions of GPK, the skill to perceive and the skill to interpret on the different indicators of opportunities to learn yielded only small effect sizes and only two predictors were statistically significant if the others were controlled for: *Knowledge* is predicted by the initial teacher education grade assigned after completion of the practical phase (second state exam[3] only, β=−.21) and *the skill to interpret* is predicted

[3] Reversed scale, i.e. the higher the grade, the worse the result: 1.0=very good, 2.0=good, 3.0=fair, 4.0=barely passed.

Table 1 Findings from multiple regression analysis to explain differences in early career mathematics teachers' general pedagogical (GP) knowledge and skills

	GP knowledge	GP skill to perceive	GP skill to interpret
Control variables			
Background variables			
Age	–.16	.04	–.01
Sex (1=male, 0=female)	–.02	.01	–.11
Predictors			
Teacher education variables			
Grade (1st state exam)	–.01	–.01	–.01
Grade (2nd state exam)	–.21*	.04	–.07
Teaching type (1=gymnasium, 0=other)	.03	–.03	.15
Teaching experience variables			
Teaching time	.03	.02	.22*
Teaching experience besides math in secondary level	.01	.13	.15
Teaching status (full time)	–.01	.08	.10
Working conditions variables			
Task demand: challenged by students	–.09	–.08	–.14
Task return: appraisal	.00	.10	.11
Variance explained in dependent variable	8.2 %	4.3 %	10.0 %

n=278

*$p \leq .05$

by teaching time compared to other responsibilities of teachers (i.e. the amount spent on preparing and conducting teaching in relation to the overall time spent on the job, β=.22). These results mean that the better the practical exam, the higher a teacher scores on the GPK test; and the larger the relative amount of time spent on teaching activities, the higher a teacher scores in the measurement of his or her skill to interpret classroom situations presented via video. Surprisingly and in contrast to our hypotheses, the skill to perceive cannot be predicted by any of the predictors included in this regression model. One possible explanation for this could be that the skill to perceive as measured by our instrument requires a level of expertise including chunks of perception that need to be developed within teaching activities ("in situ"). These chunks of perception probably will be developed in the further course of the development of expertise, but have not been developed by our target group of early career teachers yet.

In addition to these few significant effects, non-significant regression coefficients of at least small effect size ($|\beta| \geq .1$) explained some variance of the skill to interpret whose variance can be best explained (10 %). If the other predictors are controlled for, early career teachers qualified to teach mathematics at the *Gymnasium* slightly outperform (.15) those teachers who are only qualified to teach mathematics in grades 1 or 5 through 10 (elementary and middle school or middle school only). Teachers with teaching experience on other subjects than mathematics (.15) and who work full time (.10) showed slightly better test results than those without. Regarding working

conditions, teachers perceiving high level of task demand (−.14)[4], i.e. who were more often challenged by their students, and who received stronger appraisal by the principals or colleagues of their school (.11) scored slightly higher than those who felt less challenged or less appraised. Although these coefficients are not statistically significant on the 5 % level, it can be pointed out that the directions correspond with our hypotheses.

Person-Centred Analysis: Latent-Class Analysis of Competence Profiles

We conducted latent-class analysis (LCA) to investigate general pedagogical competence profiles using the two facets GPK and the general pedagogical skill to interpret classroom situations. We focussed on these two facets only since findings from the multiple regression analysis in "Variable-centred analysis: Regression of General Pedagogical Facets on Predictors" showed that the skill to perceive could not substantially be explained by the predictors.

Five models were estimated, each specifying another number of latent classes. Based on the information criteria, especially based on the BIC, which has been proven in simulation studies to be most robust (Nylund, Asparouhov & Muthén, 2007), the model specifying two latent classes revealed the smallest estimate followed by the model specifying three latent classes. In contrast, the AIC pointed to the opposite result, namely that the solution with three latent classes had a better model fit than the model specifying two classes only. In any case, the model to be chosen would be the two- or the three-class solution, since a four- or five-class solution would distinguish between more subgroups without any improvement of the BIC value, which indicates that the additional information gained is scarce. Since the model specifying three classes provided more information on the competence profile, we finally decided to choose that model.

An additional examination of the step parameters showed that the item steps were sorted as expected. Thus, the model does not include a class that deviates from the assumed order. Examining the entropy values and the average latent class probabilities for the most likely latent class pattern, we found that entropy=.63 was still in an acceptable range, whereas, even more important, average probabilities varied between .80 and .88, indicating good reliability of classification.

Figure 1 shows the z-scores of the different general pedagogical facets for each class. Although we had included GPK and the skill to interpret in the LCA only, Fig. 1 also displays the skill to perceive to present a complete picture of the three competence profiles. As can be seen, scores of GPK and the skill to interpret increase over the three profiles (therefore labeled as "low", "medium" and "high"), whereas the skill to perceive does not vary to that extent. Early career teachers assigned to the high-level profile demonstrate a higher level in the skill to perceive than the other two groups, but this difference is only significant when compared with the medium-level profile due to small group size of the low-level profile.

[4] The higher a teacher scores on that scale, the lower he or she perceives classroom management, disruptive student behavior etc. as a problem.

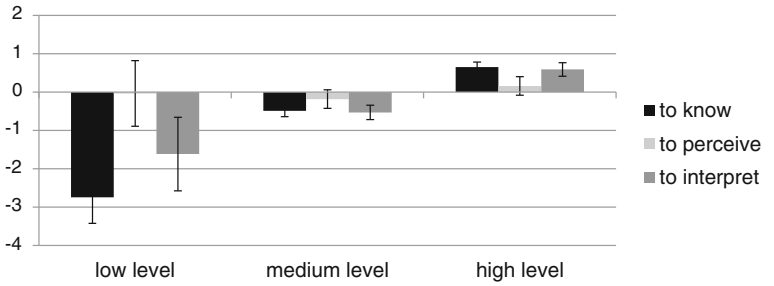

Fig. 1 Results from latent-class analysis (three-class solution)

Regression of Competence Profiles on Predictors

As a third step of data analysis, we conducted again regression analysis, but in contrast to our initial multiple regression analysis, we focussed on the profiles now to summarise the information associated with early career teachers' pedagogical competence. Thus, regression analysis was done with the class membership as dependent variable. Low level was coded with 0, medium level with 1 and high level with 2. All predictors used in our first analysis ("Variable-Centred Analysis: Regression of General Pedagogical Facets on Predictors") were included again as independent variables.

Findings in Table 2 show that, again, teaching time is a significant predictor (β=.22), thus underlining our hypothesis that pedagogical competence depends on the amount of time having spent on teaching activities. By contrast, grades are less

Table 2 Findings from regressions of general pedagogical competence profiles

	Competence profile (0=low, 1=medium, 2=high)
Control variables	
Background variables	
Age	–.19*
Sex (1=male, 0=female)	–.10
Predictors	
Teacher education variables	
Grade (1st state exam)	–.09
Grade (2nd state exam)	–.15
Teaching type (1="gymnasium", 0=other)	.05
Teaching experience variables	
Teaching time	.22*
Teaching experience besides math on secondary level	.00
Teaching status (full time)	.09
Working conditions variables	
Task demand: challenged by students	–.32*
Task return: appraisal	.06
Variance explained in dependent variable	20.7 %

n=278

*p<.05

important. None of the two state exams is statistically significant, although the grade of the second one is of small effect size and corresponds to the hypothesised direction (β= −.15). The strongest predictor in this analysis is the workplace conditions. The extent to which early career teachers feel challenged by their teaching tasks to motivate students and to master disruptive behaviour of students is statistically significant, of medium effect size, and negative (β=−.32). Thus, the stronger a competence profile, the more likely he or she reports to be challenged by his or her students regarding motivating and managing disruptive behaviour.

Discussion

Summary and Discussion of Results

In this article, we examined facets of general pedagogical competence of early career mathematics teachers, and these were general pedagogical knowledge (GPK) and the skills to perceive and interpret classroom situations, asking how they were associated with characteristics of teacher education, teaching experience, and workplace conditions. GPK was assessed via a digitalised paper-pencil test, while early career teachers' ability to perceive and interpret classroom situations was assessed via video-vignettes. Data from a follow-up study of TEDS-M Germany carried out in 2012 were used, containing a sample of 278 early career middle school teachers of mathematics.

The first step of our analysis ("Variable-Centred Analysis: Regression of General Pedagogical Facets on Predictors") showed teachers' knowledge could be predicted by teacher education grades (H1a) and teachers' skill to interpret classroom situations presented by videos could be predicted by their relative amount of time spent on teaching (H1d). Other predictors (teaching type, teaching status, experience to teach other subjects than mathematics, and task reward) did not turn out to be statistically significant, thus not supporting our hypotheses (H1b, H1c, H1e and H1f). However, one should not draw the too-far reaching conclusion that these factors would not matter regarding pedagogical knowledge and skills. All correlations were in the expected direction (although low) and some of them missed only slightly the threshold of significance.

The second step of our data analysis ("Person-Centred Analysis: Latent-Class Analysis of Competence Profiles") led us identify teachers with different general pedagogical competence profiles via LCA (H2a), while the third step ("Regression of competence profiles on predictors") provided evidence that these profiles can be predicted as hypothesised (H2b). Besides teaching experience, the competence profiles were associated with generic teaching challenges (motivating students, disruptive student behaviour) as perceived by the teachers. The higher the challenges were, the stronger the profile in terms of pedagogical knowledge and skills was. This finding let us assume that (although not that clearly hypothesised a priori, see "Research Questions and Hypotheses"), teachers who perceive a high demand of teaching tasks either do this due to their higher knowledge and skills or they cannot master these demands by just applying routines which stimulates their deliberate practice, which finally results in higher test scores. These are assumptions to be further discussed in the following, and future studies are needed to examine this assumption thoroughly.

Our results continue research on how initial teacher education provides OTL that may help pre-service teachers to acquire GPK (König, 2013) insofar as teacher education grades predict early career teachers' knowledge. In addition, the skill to interpret classroom situations seems to be determined by teaching experience. Because experience as such has been scrutinised with regards to its effect on the acquisition of knowledge (Bereiter & Scardamalia, 1993), it is interesting to see from our findings that, more specifically, it is the *relative* amount of teaching rather than a general indicator such as teaching status (working full time) which has a statistically significant effect. From that, we infer that the effect of teaching experience results mainly from careful planning of and reflecting on teaching. This finding can be further linked with the discourse on teacher reflection.

Following the notion of "reflection-in-action" proposed by Schön (1983), professional teachers are expected to identify problems in the classroom ("reflection-on-action") and be able to approach them in a way that helps them to solve the problems in the respective context, whereas it is not regarded sufficient to rely on recipe-like routines not adapted to the specific classroom situation. According to Schön (1983), these types of teacher reflection should help the teacher to build up his or her professional knowledge.

Our findings let us assume that if a teacher gives high priority to teaching then this reflects a teacher's deliberate efforts to improve his or her teaching, which finally results in the improvement of the skill to interpret classroom situations. Moreover, teachers who report that they are highly challenged may also be more strongly required to reflect on their teaching, which would finally result in higher test scores, whereas teachers who perceive no classroom management problems and who do not see the necessity to additionally motivate their students will less be urged to put effort into the improvement of their teaching.

The skill to perceive classroom situations accurately and holistically could not be related to any of the characteristics of teacher education, teaching experience, and workplace conditions, although a precise perception of classroom events is considered to be part of teacher expertise. One possible substantive explanation would be that the knowledge needed in this skill assessment has to be sufficiently internalised and consolidated on an implicit level. It serves then as an automatically activated schema and can be used without conscious effort and verbal explication. It is different from knowledge that can be explicated and, therefore, as previous findings have shown (König et al., 2014), correlates neither with the skill to interpret nor with the declarative pedagogical knowledge. This might be due to the level of expertise not yet reached by our target group of early career teachers. Moreover, we cannot rule out a methodological explanation in that reliability problems may overshadow such associations.

Conclusions and Limitations

Our analysis provides in-depth analysis on the different facets of teacher competence related to the field of general pedagogy. Findings on the influence of various factors from teacher education, teaching experience and working conditions of teachers were presented to explain what contributes to early career teachers' development of pedagogical competence. To sum up, the findings revealed once more that professional teacher competence is a multidimensional construct rather than a single homogenous

ability (König et al., 2014) and that its facets may be influenced in different ways. Initial teacher education, for example, stands for a different kind of OTL than teaching experience or working conditions of early career teachers. Seen from a general perspective, one may conclude from our study that deliberate practice is an important activity that fosters early career teachers' pedagogical competence. Our findings can be interpreted using the discourse on teacher reflection and its effects on acquisition of teacher knowledge (Schön, 1983; Dunn & Shriner, 1999; Hart et al., 2011).

Limitations of the study have to be discussed. As a follow-up study of TEDS-M, our sample is a positive selection of the TEDS-M participants. The teachers sampled in this study consist of those who at the end of their initial teacher education showed that GPK scores significantly better than TEDS-M average (see, for details, König et al., 2014). This implies a reduction of variance in GPK test scores. Since variance contributes to the strength of correlations, we cannot exclude that some of the relatively low predictors in our regression analyses may represent an underestimation. At the same time, we cannot conclude with certainty that our findings can be generalised across the full range of GPK in TEDS-M although correlations typically are fairly linear.

Furthermore, our study is only based on data from one occasion of measurement. Future research should therefore make efforts to analyse the influence of teachers' deliberate practice on the development of their pedagogical competence using longitudinal data. It would be interesting to examine how the predictors included in our study influence gain or loss of in-service teachers' pedagogical knowledge and skills. From that, it will be possible to draw more general conclusions on the issue we have started to work on, giving more profound insight into the important question whether teacher education, teaching experience, and working conditions make a difference.

References

Anderson, L. W. & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York, NY: Longman.

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.

Bereiter, C. & Scardamalia, M. (1993). *Surpassing ourselves*. Chicago, IL: Open Court.

Berliner, D. C. (1992). The nature of expertise in teaching. In F. K. Oser, A. Dick & J.-L. Patry (Eds.), *Effective and responsible teaching (Chap. 15)* (pp. 227–248). San Francisco, CA: Jossey-Bass.

Berliner, D. C. (2001). Learning about and learning from expert teachers. *International Journal of Educational Research, 35*(5), 463–482.

Berliner, D. C. (2004). Describing the behavior and documenting the accomplishments of expert teachers. *Bulletin of Science, Technology and Society, 24*(3), 200–212.

Blömeke, S. & Klein, P. (2013). When is a school environment perceived as supportive by beginning mathematics teachers? Effects of leadership, trust, autonomy and appraisal on teaching quality. *International Journal of Science and Mathematics Education, 11*, 1029–1048.

Blömeke, S., Gustafsson, J.-E. & Shavelson, R. (2015). Beyond dichotomies: Competence viewed as a continuum. *Zeitschrift für Psychologie* (in press).

Bromme, R. (1992). *Der Lehrer als Experte: zur Psychologie des professionellen Wissens [The teacher as expert: The psychology of professional knowledge]*. Bern, Switzerland: Huber.

Bromme, R. (2001). Teacher expertise. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences* (pp. 15459–15465). Amsterdam, The Netherlands: Elsevier.

Clausen, M., Reusser, K. & Klieme, E. (2003). Unterrichtsqualität auf der Basis hoch-inferenter Unterrichtsbeurteilungen: Ein Vergleich zwischen Deutschland und der deutschsprachigen Schweiz. *Unterrichtswissenschaft, 31*, 122–141.

Dehoney, J. (1995). Cognitive task analysis: Implications for the theory and practice of instructional design. Proceedings of the Annual National Convention of the Association for Educational Communications and Technology (AECT). Retrieved from ERIC database. (ED383294).

Dunn, T. G. & Shriner, C. (1999). Deliberate practice in teaching: What teachers do for self-improvement. *Teaching and Teacher Education, 15*(6), 631–651.

Ericsson, K. A., Krampe, R. T. & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*, 363–406.

Gruber, H. & Rehrl, M. (2005). *Praktikum statt Theorie? Eine Analyse relevanten Wissens zum Aufbau pädagogischer Handlungskompetenz.* [Practicum instead of Theory? Analysis of relevant Knowledge for the Acquisition of Pedagogical Performance]. Research report 15 of the Pedagogical Institute, University of Regensburg, Germany.

Hackl, B. (2004). Explizites und implizites Wissen. Menschliches Handeln im Spannungsfeld von Intentionalität, Rationalität und praktischem Können. In B. Hackl & G. H. Neuweg (Eds.), *Zur Professionalisierung pädagogischen Handelns* (pp. 69–112). Münster, Germany.

Hart, L. C., Alston, A. S. & Murata, A. (Eds.). (2011). *Lesson study research and practice in mathematics education*. New York, NY: Springer.

Hill, H. C. (2010). The nature and predictors of elementary teachers' mathematical knowledge for teaching. *Journal for Research in Mathematics Education, 41*(5), 513–545.

Hobfoll, S. E. & Freedy, J. (1993). Conservation of resources: A general stress theory applied to burnout. In W. B. Schaufeli, C. Maslach & T. Marek (Eds.), *Professional burnout: Recent developments in theory and research* (pp. 115–133). Philadelphia: Taylor & Francis.

Kaiser, G., Benthien, J., Döhrmann, M., König, J. & Blömeke, S. (2013). Expert ratings as an instrument for validating results of video-based testing. In A. Lindmeier & A. Heinze (Hrsg.). *Proceedings of the 37th conference of the international group for the psychology of mathematics education*. (Bd. 1). Kiel, Germany: PME.

Keller-Schneider, M. & Hericks, U. (2011). Forschungen zum Berufseinstieg. Übergang von der Ausbildung in den Beruf [Research on Entering the Profession. Transition from Training into Profession]. In E. Terhart, H. Bennewitz & M. Rothland (Eds.), *Handbuch der Forschung zum Lehrerberuf [Handbook on the Research of Teachers]* (pp. 296–313). Münster, Germany: Waxmann.

Kersting, N. B., Givvin, K. B., Thompson, B. J., Santagata, R. & Stigler, J. W. (2012). Measuring usable knowledge teachers' analyses of mathematics classroom videos predict teaching quality and student learning. *American Educational Research Journal, 49*(3), 568–589.

Klein, G. A. & Hoffman, R. R. (1993). Seeing the invisible: Perceptual–cognitive aspects of expertise. In M. Rabinowitz (Ed.), *Cognitive science foundations of instruction* (pp. 203–226). Hillsdale, MI: Erlbaum.

König, J. (2013). First comes the theory, then the practice? On the acquisition of general pedagogical knowledge during initial teacher education. *International Journal of Science and Mathematics Education, 11*(4), 999–1028.

Konig, J. (2014). Designing an international instrument to assess teachers' general pedagogical knowledge (GPK): Review of studies, considerations, and recommendations. Technical paper prepared for the ITEL project. Paris: OECD.

König, J. & Blömeke, S. (2012). Future teachers' general pedagogical knowledge from a comparative perspective. Does school experience matter? ZDM - *The International Journal on Mathematics Education, 44*(3), 341–354.

König, J. & Lebens, M. (2012). Classroom management expertise (CME) von Lehrkräften messen: Überlegungen zur Testung mithilfe von Videovignetten und erste empirische Befunde. *Lehrerbildung auf dem Prüfstand, 5*(1), 3–29.

König, J. & Blömeke, S. (2013). Preparing teachers of mathematics in Germany. In J. Schwille, L. Ingvarson & R. Holdgreve-Resendez (Eds.), *TEDS-M encyclopaedia. A guide to teacher education context, structure and quality assurance in 17 countries. Findings from the IEA teacher education and development study in mathematics (TEDS-M)* (pp. 100–115). Amsterdam: IEA.

König, J., Blömeke, S., Paine, L., Schmidt, B. & Hsieh, F-J. (2011). General pedagogical knowledge of future middle school teachers. On the complex ecology of teacher education in the United States, Germany, and Taiwan. *Journal of Teacher Education, 62*(2), 188–201.

König, J., Blömeke, S., Klein, P., Suhl, U., Busse, A. & Kaiser, G. (2014). Is teachers' general pedagogical knowledge a premise for noticing and interpreting classroom situations? A video-based assessment approach. *Teaching and Teacher Education, 38*, 76–88.

Kouzes, J. M. & Posner, B. Z. (1999). *Encouraging the heart: A leader's guide to rewarding and recognizing others*. San Francisco, CA: Jossey-Bass.

Magidson, J. & Vermunt, J. K. (2004). Latent class analysis. In D. Kaplan (Ed.), *The Sage handbook of quantitative methodology for the social sciences (Chap. 10)* (pp. 175–198). Thousand Oaks, CA: Sage.

McCormack, A., Gore, J. & Thomas, K. (2006). Early career teacher professional learning. *Asia-Pacific Journal of Teacher Education, 34*(1), 95–113.

Muthén, B. O. & Muthén, L. K. (1998–2006). Mplus (Version 4.2) [Computer software]. Los Angeles, CA.

Nylund, K. L., Asparouhov, T. & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal, 14*, 535–569.

Putnam, R. T. (1987). Structuring and adjusting content for students: A study of live and simulated tutoring of addition. *American Educational Research Journal, 24*, 13–48.

Rost, J. & Georg, W. (1991). Alternative Skalierungsmöglichkeiten zur klassischen Testtheorie am Beispiel der Skala, Jugendzentrismus". *ZA-information, 28*, 52–74.

Sabers, D. S., Cushing, K. S. & Berliner, D. C. (1991). Differences among teachers in a task characterized by simultaneity, multidimensionality, and immediacy. *American Educational Research Journal, 28*, 63–88.

Schoenfeld, A. H. (2011). Reflections on teacher expertise. In Y. Li & G. Kaiser (Eds.), *Expertise in mathematics instruction* (pp. 327–341). US: Springer.

Schön, D.A. (1983). *The reflective practitioner—how professionals think in action*. New York, NY, Basis Books.

Shavelson, R. J. (2010). On the measurement of competency. *Empirical Research in Vocational Education and Training, 2*(1), 82–103.

Sherin, M., Jacobs, V. & Philipp, R. (Eds.). (2011). *Mathematics teacher noticing: Seeing through teachers' eyes*. Routledge.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Research, 57*, 1–22.

Tatto, M. T., Schwille, J., Senk, S., Ingvarson, L., Peck, R. & Rowley, G. (2008). *Teacher Education and Development Study in Mathematics (TEDS-M): Conceptual framework*. East Lansing: Michigan State University.

Valentine, J., Clark, D., Hackmann, D. & Petzko, V. (2004). *Leadership for highly successful middle level schools. Volume II: A national study of leadership in middle level schools*. Reston, VA: National Association of Secondary School Principals.

van Es, E. A. & Sherin, M. G. (2002). Learning to notice: Scaffolding new teachers' interpretations of classroom interactions. *Journal of Technology and Teacher Education, 10*(4), 571–596.

von Eye, A. & Bergman, L. R. (2003). Research strategies in developmental psychopathology: Dimensional identity and the person-oriented approach. *Development and Psychopathology, 15*(3), 553–580.

Weinert, F. E. (2001). A concept of competence: A conceptual clarification. In D. S. Rychen & L. H. Salganik (Eds.), *Defining and selecting key competencies* (pp. 45–65). Göttingen, Germany: Hogrefe.

Wu, M. L., Adams, R. J. & Wilson, M. R. (1997). ConQuest: Multi-aspect test software (computer program). Camberwell: Australian Council for Educational Research.