JOHANNES KÖNIG

FIRST COMES THE THEORY, THEN THE PRACTICE? ON THE ACQUISITION OF GENERAL PEDAGOGICAL KNOWLEDGE DURING INITIAL TEACHER EDUCATION

Received: 13 September 2012; Accepted: 08 April 2013

ABSTRACT. Teacher education systems worldwide are confronted with the essential question of how to foster both future teachers' theoretical and practical knowledge and to adequately enable future teachers to connect their theoretical and practical knowledge for teaching. This article investigates how future teachers acquire general pedagogical knowledge (GPK) as a central component of teacher knowledge during initial teacher education, exemplified by pre-service teachers in Germany, where initial teacher education is divided into a first phase with a heavy focus on theoretical, academic study, and a second phase where future teachers learn how to apply their theoretical knowledge in the classroom. Data from teacher knowledge studies Teacher Education and Development Study in Mathematics and *Längsschnittliche Erhebung pädagogischer Kompetenzen von Lehramtsstudierenden*/Longitudinal Survey of Student Teachers' Pedagogical Competencies are used to compare future teachers' GPK at different teacher education stages (the beginning, after 2 years, and end of training). Findings show the more advanced future teachers are in the course of their initial teacher education, the better they perform in the test measuring GPK. When analyzing subscales of the test measuring cognitive dimensions of GPK, as would be expected declarative-conceptual knowledge (measured by cognitive dimensions "recall" and "understand/analyze") was gained predominantly during the theoretical study (first phase), whereas future teachers who had additionally passed through the practical second phase performed much better on the practical knowledge test subscale (measured by the cognitive dimension "generate"). Research findings are discussed with regard to the development of teacher expertise during initial teacher education, and recommendations for future policy directions with respect to teacher education are given.

KEY WORDS: assessment, general pedagogical knowledge, opportunities to learn, pre-service teachers, procedural knowledge, teacher education, teacher expertise

Introduction

Current research on teacher knowledge usually distinguishes three domains of professional knowledge for teaching (Baumert & Kunter, 2006; Bromme, 1997; Grossman & Richert, 1988; Shulman, 1986, 1987): content knowledge (CK), pedagogical content knowledge (PCK), and general pedagogical knowledge (GPK). Although there is a growing body of research in the field of teacher knowledge, there is a special focus on

International Journal of Science and Mathematics Education (2013) 11: 999–1028

the two content-related domains, mostly as exemplified by mathematics teachers (e.g. Ball, Thames & Phelps, 2008; Schmidt, Cogan, Houang, 2011; Kunter, Baumert, Blum, Klusmann, Krauss, Neubrand, 2011). Thus, regarding the domain of GPK, there are still open questions such as: What is exactly meant by the term GPK and what does this knowledge domain incorporate? How do future teachers acquire GPK during initial teacher education and how do opportunities to learn in the field of general pedagogy, as well as field experience support future teachers' acquisition of GPK?

These questions need to be investigated by empirical research, for at least three reasons. First, professional teacher knowledge is multidimensional (Baumert & Kunter, 2006; Fenstermacher, 1994), involving integrating elements of general pedagogical and subject-related knowledge and taking into account their being of a different quality. Secondly, opportunities to learn during initial teacher education are provided not only by subject-related disciplines but also by disciplines related to general pedagogy. Initial teacher education intends to support future teachers' acquisition of professional knowledge for teaching. Regarding GPK, however, little is known about how future teachers acquire knowledge of this domain and thus how they may be supported to progress from the stage of teacher "novices" to "advanced beginners" (Berliner, 2001, 2004). Thirdly, discussions about the reform of teacher education are often dominated more by evaluative than evidence-based statements (Ball et al., 2008; König & Blömeke, in press). Without any information about the acquisition of professional knowledge during initial teacher education resulting from empirical testing, such discussions have their limits in the process of improving teacher education (Larcher & Oelkers, 2004).

This article tries to work on this problem. It reports on the theoretical framework of a standardized paper-and-pencil test measuring GPK that was developed in the context of the Teacher Education and Development Study in Mathematics (TEDS-M) carried out under the supervision of the International Association for the Evaluation of Educational Achievement (IEA).[1] TEDS-M was a comparative study of teacher education and the first IEA study on tertiary education, as well as the first international large-scale assessment of future teachers that worked with representative samples (Tatto, Schwille, Senk, Ingvarson, Rowley, Peck, Bankov et al., 2012). The TEDS-M target population consisted of mathematics teachers for elementary and middle schools in their final year of teacher education. Data were collected in 2008. A central component of TEDS-M was to measure the professional knowledge of future teachers. However, the common international questionnaire only measured future teachers' mathematics content knowledge (MCK) and mathematics pedagogical

content knowledge (MPCK). Three participating countries—the USA, Germany, and Taiwan—therefore decided to participate in a national option measuring future teachers' GPK. A framework specifying significant elements of GPK that are relevant to describe future teachers' knowledge acquisition during initial teacher education and a test instrument were developed under the leadership of the German TEDS-M team (for details see König & Blömeke, 2009; König, Blömeke, Paine, Schmidt & Hsieh, 2011). In Germany, besides participating in this national option, additional data were collected by another study (LEK–*Längsschnittliche Erhebung pädagogischer Kompetenzen von Lehramtsstudierenden*/Longitudinal Survey of Student Teachers' Pedagogical Competencies) that sampled the first-year pre-service teacher cohort in 2008 (König & Seifert, 2012).[2] They were tested twice, at the very start of their teacher training and two years after when they had been trained at university for four semesters. The following analysis makes use of both the TEDS-M and the LEK data to compare general pedagogical knowledge assessments during the course of teacher education allowing an in-depth analysis of the test measuring GPK. We focus on the question of how the different quality of GPK as mirrored by cognitive dimensions will be relevant to describe future teachers' acquisition of knowledge in the course of teacher education.

Conceptual Framework

The GPK of Future Teachers

According to Shulman (1987, p. 8), GPK involves "broad principles and strategies of classroom management and organisation that appear to transcend subject matter" as well as knowledge about learners and learning, assessment, and educational contexts and purposes. Similarly, and extending this definition, Grossman & Richert (1988, p. 54) stated that GPK "includes knowledge of theories and learning and general principles of instruction, an understanding of the various philosophies of education, general knowledge about learners, and knowledge of the principles and techniques of classroom management." Future teachers need to draw on this range of knowledge and weave it into coherent understandings and skills if they are to become competent to deal with what McDonald (1992) called the "wild triangle" that connects learner, subject matter, and teacher in the classroom.

However, because there was a lack of empirical studies on (future) teachers' GPK (Wilson & Berne, 1999), when the need to measure GPK in

the context of TEDS-M 2008 came up, many key questions were unanswered. There were virtually no studies showing how to fill these relatively broad domains of GPK so that one could develop items and actually test teachers (Baumert & Kunter, 2006). Another open question was how to discriminate GPK from MPCK. In Germany and Switzerland, some first attempts existed to measure the GPK of future teachers (Baer, Dörr, Fraefel, Kocher, Küster, Larcher et al., 2007; Schulte, 2008), but these studies were restricted to specific institutions, languages, or regions, and the data collected by these studies contained only a relatively small number of future teachers. A substantial number of studies had tried to capture GPK with self-reports of future teachers (e.g. Oser & Oelkers, 2001), but these did not include objective tests (for a critical discussion of this issue see König, Kaiser & Felbrich, 2012). Against the background of this research deficit, a theoretical framework of future teachers' GPK that could be tested empirically across countries in the context of TEDS-M had to be developed. Due to the complexity of GPK and the audience of an international survey, as well as with regard to standardized test procedures on a large scale and the target population of future teachers (and not practicing teachers), it was necessary to make certain restrictions in the definition of general pedagogy.

Following the concept of "competence" (see in general Weinert, 2001; specified for the teaching profession by Bromme, 1992, 1997, 2001), the theoretical framework of GPK developed in the context of TEDS-M is structured in a task-based way and explicitly not according to the formal structure of general pedagogy as an academic discipline. Furthermore, instruction is focused on as the core activity of teachers (Baumert & Kunter, 2006; Berliner, 2001, 2004; Bromme, 1997) serving as a heuristic to select the topics and cognitive demands of GPK. Findings from instructional research (Good & Brophy, 2007; Helmke, 2003; Slavin, 1994) and didactics (cf. Good & Brophy, 2007; Klafki, 1985; Tulodziecki, Herzig & Blömeke, 2004) were combined to conceptualise GPK for teaching, as shown in Figure 1 (for details, see König et al.,

Test dimensions	Topics covered by the test items
Structure	- structuring of learning objectives - lesson planning and structuring the lesson process - lesson evaluation
Motivation/ classroom management	- achievement motivation - strategies to motivate single students/ the whole group - strategies to prevent and counteract interferences - effective use of allocated time/ routines
Adaptivity	- strategies of differentiation - use of a wide range of teaching methods
Assessment	- assessment types and functions - evaluation criteria - teacher expectation effects

Figure 1. Dimensions and topics covered in the TEDS-M test of GPK

2011): Four dimensions of GPK are considered highly relevant with respect to the target group of future teachers. Teacher education is regarded as effective if future teachers in their last year of their training have acquired general pedagogical knowledge allowing them to prepare, structure and evaluate lessons ("structure"), to motivate and support students as well as manage the classroom ("motivation/classroom management"), to deal with heterogeneous learning groups in the classroom ("adaptivity") and to assess students ("assessment").

Cognitive Dimensions of GPK

Apart from the task-based dimensions and topics of GPK shown in Figure 1, dimensions of cognitive processes were additionally defined to describe the cognitive demands on future teachers when they respond to test items. Following Anderson and Krathwohl's elaborate and well-known model (2001), with the TEDS-M test instrument measuring GPK, three cognitive processes were distinguished which summarise the original six processes: recalling, understanding/analyzing, and generating.

1. Recalling. Future teachers have to retrieve information from long-term memory in order to respond to a test item. Test items of this type challenge future teachers to give an example for a definition, to recite elements of a phenomenon, a term, a concept, or to identify a term or a concept.
2. Understanding/analyzing. In order to respond to a test item of this type, future teachers also have to retrieve information from long-term memory, but, moreover, they have to link that information with a problem outlined by the test item. So they have to describe or explain a phenomenon or a concept; or they are asked to compare, categorize, assign, or interpret a phenomenon, a situation, or one or several general terms.
3. Generating. To respond to items of the third dimension of cognitive processes, future teachers have to generate concrete strategies concerning how they would solve a typical classroom situation problem which includes evaluating this situation. Again, retrieving information from long-term memory might be helpful, but, moreover, that knowledge has to be linked with classroom situation experience. Future teachers are asked to explicate practical knowledge that can be described as a propositional mental representation. This item type reflects the need to measure GPK that is of a situated nature (Putnam & Borko, 2000).

The definitions of cognitive processes suggest GPK is of a different quality. Furthermore, they can be related to the well-known differentiation between declarative knowledge ("knowing that…") and procedural knowl-

edge ("knowing how…"), a distinction that is typical for research on teacher knowledge (besides Anderson & Karthwohl, 2001, see Fenstermacher, 1994; Bromme, 2001). While the first two dimensions of cognitive processes (recalling and understanding/analyzing) predominantly test declarative or conceptual knowledge, items of the third dimension (generating) tend to measure procedural knowledge.

Especially, research on teacher expertise has determined that declarative and procedural knowledge contribute to the expert's performance in the classroom (Bromme, 2001). While declarative and procedural knowledge are of a specific quality, they are interdependent (Fenstermacher, 1994). Declarative knowledge is frequently regarded as a premise for procedural knowledge. While declarative knowledge is proceduralized (Anderson, 1982), knowledge, described as competencies, merges into performance on a higher level (Klieme, Avenarius, Blum, Döbrich, Gruber, Prenzel, Reiss et al., 2003).

Item Examples Measuring GPK

In TEDS-M, topics of GPK and cognitive demands made up a matrix which served as a heuristic for item development (see Figure 2). For each cell, a subset of items was developed. Several expert reviews, as well as two large pilot studies were carried out. All the experts who participated in the first item review, which aimed at selecting items for the first pilot study testing a large pool of items, were teacher educators in the field of general pedagogy. Moreover, their research had to be related to the topic of teacher knowledge, and they had to be at least PhD candidates. In contrast to this first expert review, experts that participated in the second and following reviews, which aimed at selecting items for the final test instrument according to specific criteria or at validating the test instrument, respectively, had to endow a university chair with a specialization on research about teacher knowledge. Based on these review processes and empirical findings from the two pilot studies (e.g. item parameter estimates), as well as on conceptual considerations with respect to the framework, the final item set was selected (König & Blömeke, 2010a, b, König et al., 2011).

	Recall	Understand/ analyze	Generate
Structure			
Motivation/classroom management			
Adaptivity			
Assessment			

Figure 2. Test design matrix

Three item examples (see Figures 3, 4, and 5) may illustrate the GPK test and the heuristic used to conceptualize GPK (see Figures 1 and 2). [3] 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 analyze five statements against the background of this distinction. Statement C represented an example of "intrinsic motivation," whereas A, B, D, and E were examples for "extrinsic motivation."

The second and third item examples are open-response items (see Figures 4 and 5). Regarding the item shown in Figure 4, future teachers were asked to support another future teacher and evaluate their lesson. This is a typical challenge during a peer-led teacher education practicum, but practicing teachers are also regularly required to analyze 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. Similarly, when responding to the test item shown in Figure 5, future teachers had to "generate" different methods that would be useful when providing feedback to their students at the end of a lesson.

For the open-response items, coding rubrics were developed and reviewed by experts on teacher education. First, coding instructions were developed in a complex and extensive interplay of deductive (from our theoretical framework) and inductive approaches (from empirical teacher responses). In a pilot phase, codes from several independent raters were discussed in detail, and coding instructions were carefully revised and expanded. The result was then reviewed by experts. Thus, the coding

Which of the following cases represents an example of intrinsic motivation, and which represents an example of extrinsic motivation?

Check one box in each row.

A student learns before a test in mathematics, because he/she...		**intrinsic motivation**	**extrinsic motivation**
A.	expects a reward for a good grade.	☐	☐
B.	wants to avoid the consequences of a bad grade.	☐	☐
C.	is interested in problems of mathematics.	☐	☐
D.	does not want to disappoint his/her parents.	☐	☐
E.	wants to maintain his/her relative rank in the class.	☐	☐

Figure 3. Item example 1 for GPK about "motivation" and "analyze"

Imagine you are helping a future teacher to evaluate her lesson because she has never done this before.
To help her adequately analyze her lesson, what question would you ask?
Formulate ten essential questions and write them down.

Figure 4. Item example 2 for GPK about "structure" and "generate"

manual is theoretically based, as well as data-based. The codes were intended to be low-inferent, i.e. every response was coded with the least possible amount of inferences by the raters.

In TEDS-M, all questionnaires were coded on the basis of the coding manual. Two raters coded the answers independently of one another. As a measure of consensus and internal consistency, Cohen's Kappa was estimated (Jonsson & Svingby, 2007). It ranges from 0.80 to 0.99 with an average of $M = 0.91$ (SD = 0.07). This can be regarded as a good result. If conformity of raters was lacking, an agreement between the two raters was obtained in collective discussion, calling on a third rater if necessary.

After having established reliable *coding* schemes, *scoring* strategies for complex open-response items such as the two shown in Figures 4 and 5 were developed in order to decide which codes could be rewarded and which could not because they did not seem to be appropriate. Again, experts had to agree on appropriate answers which would sufficiently reflect expected outcomes of teacher education. Illustrating this strategy with the test item shown in Figure 4, codes were scored as appropriate if they addressed four criteria: "context" of the lesson (e.g. prior knowledge of students), "input" (e.g. objectives of the lesson), "process" (e.g. teaching methods used), and "output" of the lesson (e.g. student achievement). The extract of an original answer given by a future US teacher in the TEDS-M survey (see Figure 6) is a good example of these four criteria. For the test item in Figure 5, codes were scored as appropriate if they addressed the three criteria: "teacher-centered" (e.g. oral inquiry by the teacher), "interactive" (e.g. discussion of work results), and "student-centered" (e.g. self-evaluation by students).

Initial Teacher Education and the Acquisition of GPK

Initial teacher education programs aim at preparing students to become well-qualified teachers. Among other goals that might be pursued, such programs intend that future teachers acquire professional knowledge. Thus, subject-related, but also general pedgagogical opportunities to learn

At the end of a lesson, how can students receive useful feedback about their learning during the lesson?
Give three different methods you find exceptionally useful.

Figure 5. Item example 3 for GPK about "assessment" and "generate"

1) *Do your students have prior knowledge about the subject?*
2) *What are your objectives?*
3) *Are the students working individually or in groups?*
…
10) *Have your students gained the knowledge from the lesson?*

Figure 6. US future teacher's response to item example 2

are provided by teacher education institutions (Clift & Brady, 2005; Schmidt, Tatto, Bankov, Blömeke, Cedillo, Cogan, Han et al., 2007). While courses in the academic setting often primarily aim at the acquisition of theoretical knowledge, in-school opportunities to learn give future teachers the chance to connect their knowledge to practical situations in the classroom. This approach is discussed as the ideal way to achieve the professionalism of future teachers (Dann, 2000; Kolbe & Combe, 2004), since then future teachers start to acquire and reflect practical knowledge.

With regard to GPK, as defined for the measurement in TEDS-M, such practical experiences whereby future teachers have the chance to teach students in the classroom should turn out to be particularly important. Future teachers should then be forced to reflect on tasks such as structuring lessons, dealing with heterogeneity, or motivating students, and thus to activate the GPK that they might have already previously acquired in the academic context. Presumably, while making use of GPK in such situations, future teachers become increasingly flexible in how to apply their knowledge, starting to transform declarative knowledge into procedural knowledge (Anderson, 1982; Hatano & Inagaki, 1986; Berliner, 2001, 2004; Gruber & Rehrl, 2005; König, 2010). Future teachers are challenged to integrate "propositional knowledge with practical knowledge in the field" (Clift & Brady, 2005, p. 332).

However, empirical research on the effectiveness of teacher education does not tell us precisely how different kinds of teacher education program opportunities to learn are related to GPK. This is mainly due to the lack of studies measuring GPK, or performance of pre-service teachers in general (König, 2012; König & Seifert, 2012). So, for instance, Clift & Brady (2005, p. 332) conclude that "we know little about how practice plays out in school settings and what consequence recommended practice has for students."

Initial Teacher Education in Germany

In being compared internationally, Germany is a country with a very specific teacher education structure since it appears to offer teacher education programs that are spread over two phases, a theoretical and a practical one (König & Blömeke, in press), that are offered by two

completely different institutions. Future teachers begin their preparation in one of the German universities with programs that emphasize academic theoretical study. This assures a relatively advanced level of academic preparation for all future teachers. The first phase also contains a great deal of required education coursework with a heavy emphasis on theory. In total, when TEDS-M was carried out, 74 universities provided the first phase of teacher education in Germany. Future teachers finish university with a degree (First State Examination) which is equivalent to a Masters or Diploma, thus also allowing them to apply for jobs others than teaching or starting a dissertation leading to a doctoral degree. The completion of the first phase and the first state examination certificate are the general requirement for entry into the second phase. By contrast, most of the practical preparation is provided in the second phase in special, generally small institutions operated by state governments and known as "Studienseminare". Future teachers must work part-time at schools and attend courses in general pedagogy (*Hauptseminar*) and subject-related pedagogy (*Fachseminar*). The second phase ends with the Second State Examination consisting of a practical part including at least two lessons performed in two different subjects and an oral examination. The examination includes the requirement to write an essay (*Zweite Staatsexamensarbeit*) where future teachers describe the planning, teaching, and analysis of a sequence of lessons they have taught. Future teachers are assessed by teacher educators whose courses they have to attend during the second phase, and they are mentored by one or two teachers at a school.

With regard to general pedagogy opportunities to learn, there is a general tendency that theoretical study is done at university, while learning to apply that knowledge is subject to the second phase (Terhart, 2003; Messner, 2004). From this, it can be inferred that future teachers predominantly acquire theoretical-formal knowledge (Fenstermacher, 1994) during the first phase of teacher education, while they mainly acquire procedural knowledge after having finished university courses (Terhart, 1993). Practical opportunities to learn provided by the first phase are of short duration only, usually between 2 and 4 weeks (König & Blömeke, in press), so that they are only regarded as isolated opportunities to learn to foster future teachers' performance in class (Hascher, 2006; König, 2012).

However, the sequence of theoretical training first, followed by the practical transfer as a second step during initial teacher education has received approval. Advocates state that the professional performance of teachers is founded on cognitive skills helping the teachers to notice,

analyze, and reflect the complexity of classroom events (cf. Bromme, 1992, 1997; Gerstenmaier & Mandl, 2000). Teachers' professional GPK therefore does not only contain performance mastery in the classroom, but includes basic knowledge on the mastery of central professional tasks, so that teachers are also able to analyze and reflect on their performance results and to generate useful performance strategies in the classroom (Messner, 2004). To be concrete, a good example is lesson planning. Future teachers exercise lesson planning by writing down every single didactic-methodical step, and they are asked to reflect in slow motion whether each option they think of is right and what alternative options exist. This process may finally contribute to the overall success of the lesson plan.

Research Questions

Although the two phases of initial teacher education in Germany seem to stand for a virtually unique initial teacher education system structure worldwide, from a general point of view many teacher education systems are confronted with the essential question of how to provide appropriate opportunities for learning that foster both future teachers' theoretical and practical knowledge and that adequately enable future teachers to connect the theoretical and practical issues of teaching. This is why, for example, many countries have made great efforts to set up induction programs, some of which could even be compared with the second phase of initial teacher education in Germany.

Against this background, the following analysis serves to investigate the question of how future teachers acquire GPK in the course of initial teacher education, exemplified by future teachers undergoing teacher education in Germany. Two major research questions are focused on:

1. Do future teachers gain GPK in the course of their training?

Since teacher education in Germany claims to be effective and the conceptualization of the test measuring GPK claims all dimensions underlying the test instrument are curricular valid with regard to German teacher education, we assume we are able to measure the continuous knowledge gain of future teachers from the very beginning when they enter teacher education to the point of time when they would have had 2 years training by the end of their teacher education. Empirical findings from testing future teachers' GPK supporting this assumption will extend the TEDS-M international research on GPK (König et al., 2011) since TEDS-M as a cross-sectional study only informs about how future teachers

perform at the end of their training. Findings from TEDS-M do not tell us whether the future teachers' GPK tested actually results from learning processes in teacher education programs or, following Lortie's notion of the "apprenticeship of observation" (Lortie, 1975, p. 61), might have existed even before future teachers had entered teacher education as a result of the future teachers' socialisation as a student having taken place in the school and classroom they had attended for many years.

2. Do future teachers' acquisition of knowledge depend on opportunities to learn that focus on theoretical and practical issues of general pedagogy?

Since the German initial teacher education's first phase tends to foster future teachers' theoretical knowledge primarily, we expect that future teachers acquire GPK related to the first two cognitive dimensions of the test instrument (recalling and understanding/analyzing) between the first occasion of measurement (entry to teacher education) and the second occasion of measurement after 2 years. By contrast, when comparing test results between future teachers having been educated for 2 years only and future teachers in their last year of training we expect better test results, particularly with regard to the third cognitive dimension (generating) to the benefit of the latter group since they have had in-school opportunities to learn to a much larger extent, and it is the main objective of the German initial teacher education's second phase to foster future teachers' performance in a class (König & Blömeke, in press).

Methodology

Sample

For the purpose of this article, data from TEDS-M and LEK are used. The TEDS-M target population was defined as "future teachers who are in their final year of training before they are eligible to become practicing teachers of mathematics in primary and in lower secondary schools (either as generalist teachers or as mathematics specialists)" (Tatto, Schwille, Senk, Ingvarson, Peck, Rowley, 2008, p. 32). TEDS-M distinguished between the primary and the secondary level and therefore conducted two surveys, one sampling future primary teachers and another sampling future secondary teachers (Tatto et al., 2012).

With regard to the primary future teacher survey, a teacher education program was identified as preparing primary teachers if the license

included one of the grades 1 through 4. In a two-stage process, random samples were drawn from this target population in each participating country. The samples were stratified according to important teacher education features like "route" (consecutive vs. concurrent programs), "type" of program (grade span the license includes), or "region" (e.g. federal states) in order to accurately reflect the distribution of future primary school teachers' characteristics at the end of their training. In 2008, about 14,000 future primary teachers from 16 countries were surveyed, including 1,032 future teachers from Germany. Germany successfully met IEA's quality requirements as known from TIMSS or PIRLS, which included monitoring of test situations and meeting participation rates. In Germany, the 30-min GPK test was taken after the official 90-min TEDS-M future teacher survey, a design implemented with the permission of the TEDS-M International Study Center.

The LEK target population was defined as future teachers who had started teacher training in the winter term, 2008, at four universities that were deliberately selected according to the research design criteria (König & Seifert, 2012). Since at that point of time 74 universities provided pre-service teacher education in Germany (König & Blömeke, in press), it is obvious the sample of four universities is not capable of allowing generalization of results. Also, in contrast to TEDS-M, future teachers of all subjects were sampled.

However, there is a relevant intersection of the TEDS-M and LEK target group definitions since in Germany, all primary future teachers have to undergo mathematics courses. Moreover, there are two types of primary teacher education programs in Germany (König & Blömeke, in press). Future primary teachers get qualified either to teach at primary schools only (grades 1 to 4) or to teach at primary and lower secondary schools (grades 1 to 9/10). In 2008, seven federal states and their universities in Germany provided the first program type, whereas the second program type was provided by the other nine federal states and their universities. Since the four universities sampled in LEK belonged to the first group of federal states providing the first type, in the following analysis only the TEDS-M data subset is used that contains the representative future primary teacher sample following a teaching career path to teach at primary schools from grades 1 to 4 (n = 522). The relevant LEK sample includes 294 future primary teachers who participated in the first occasion of measurement (i.e. when entering teacher training) and 193 future primary teachers who participated in the second occasion of measurement. As in TEDS-M, future teachers participating in LEK were instructed by a trained test administrator and tested under observation. They were given 30 min to complete the GPK

test so that comparability of test results between TEDS-M and LEK was assured (for further details of data collection in LEK, see König & Seifert, 2012).

Test Instrument and Scaling Analysis

As laid out above, teacher tasks and cognitive demands made up a matrix which served as a heuristic for the development of the GPK items (see Figure 2). For each cell, a subset of items was developed (for details of test development, see König et al., 2011). The instrument used to measure the GPK of future primary teachers in TEDS-M as well as LEK consisted of 85 test items. These included dichotomous and partial-credit items as well as open-response (about half of the test items) and multiple-choice items. The items were fairly equally distributed across the four teacher tasks and the three cognitive demands. Following the TEDS-M test design for MCK and MPCK (for details see Tatto et al., 2008), a balanced incomplete block design (Adams & Wu, 2002; von Davier, Carstensen, von Davier, 2006) with five booklets was used so that each teacher had to respond to only two thirds of the test items.

Item Response Theory (IRT) scaling methods were used to estimate scores across the different booklets. With the procedures implemented in a software package such as ConQuest (Wu, Adams, Wilson, 1997), it is possible to create reliable achievement scores even if a person had responded to only a selection of test items if this selection was done rigorously according to a range of specific criteria.

Figure 7 shows an item-person map from the unidimensional IRT analysis that was done in TEDS-M to compare primary future teachers in Germany and the USA (König & Blömeke, 2010a, p. 268). On the left side abilities of future teachers are represented (one "X" represents 11 people), whereas on the right side the distribution of test items is shown (each of the 85 test items has a number). If the location of an item and a person match, the person has a probability of 0.5 of succeeding on that item. The higher a person is above an item on the scale, the more likely the person will succeed on the item. The lower a person is below an item on the scale, the more likely the person will not be able to solve the item successfully. The GPK test covered the ability range of the TEDS-M sample of future primary teachers from the USA and Germany reasonably well as the range of the personal abilities (left side) was well covered by the item difficulties (right side). The one-dimensional model and its results showed that it was possible to create an overall GPK test score. The reliability was good (EAP reliability 0.86).[4]

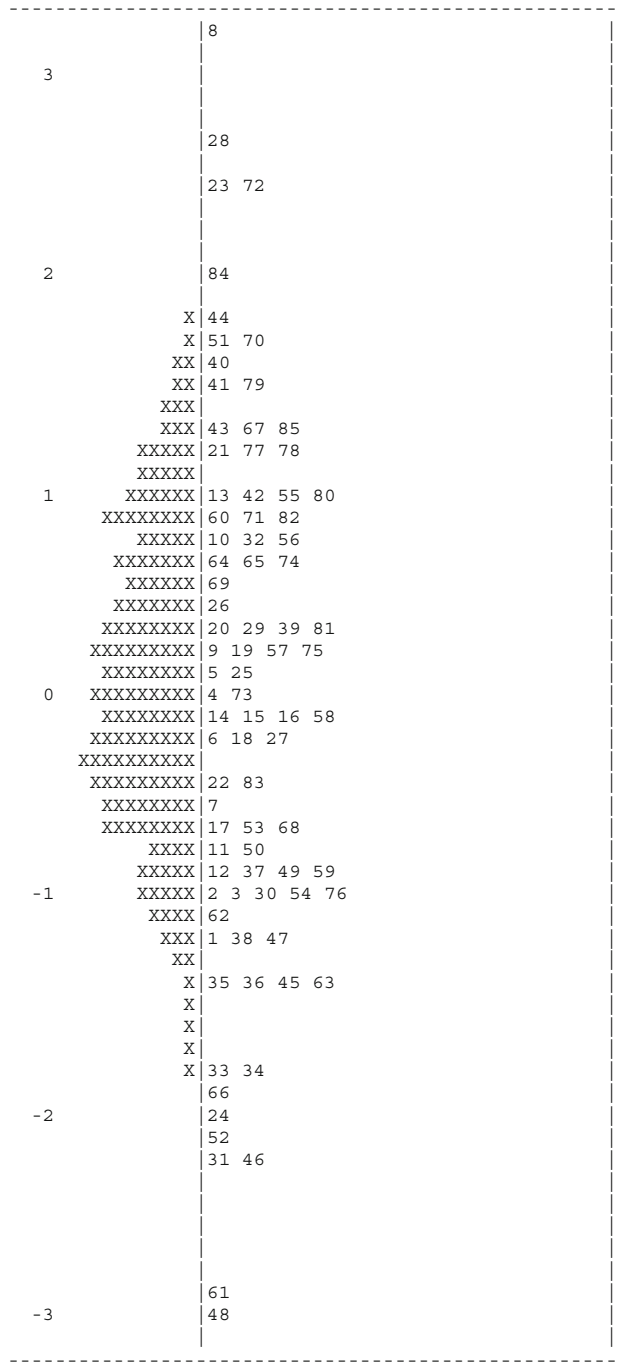

Figure 7. Item-person map of one-dimensional IRT scaling (König & Blömeke, 2010a)

In a second step, the test scores of future primary teachers who had also participated in LEK were estimated by using IRT scaling methods. To assure comparability, GPK test item parameter estimates were exported from the TEDS-M scaling analysis as is shown in Figure 7 and then used for the scaling of LEK future primary teachers' test data. [5] In the LEK data scaling analysis, we treated all observations as independent, providing a larger number of observations (n = 487) and thus effectively increasing the analytical power of the final scaling analysis (cf. Bond & Fox, 2007). The reliability of the GPK overall scale was good (EAP reliability 0.90). Finally, individual ability estimates exported from IRT modeling were transformed using the same formula applied in TEDS-M to facilitate the reading of test results (TEDS-M international mean 500, TEDS-M international standard deviation of 100 test points).

In contrast to a model in which all items measure one latent ability (see model on the left side in Figure 8), in a multidimensional IRT model test, items were scaled as hypothesized and documented in our conceptual framework according to the cognitive processes requested (recall, understand/analyze, and generate; see model on the right side in Figure 8). The reliability of each of these three subscales was acceptable for the subscales Recall and Understand/Analyze (Table 1). The reliability for the third subscale, Generate, was rather low, however, indicating that this subdimension was difficult to measure. Nonetheless, the reliability could still be regarded as sufficient in order to describe future primary school teachers' GPK in that area. Intercorrelations between subscales measuring the different cognitive processes are statistically significantly lower for the beginning teachers (LEK) compared with the future teachers in their last year of training (TEDS-M), showing that GPK really is a heterogeneous construct in the first phase of teacher education, but becomes increasingly consolidated in the second phase of teacher education (see Table 2).

Results

Results on the Overall GPK Score and the Cognitive Dimensions of GPK

First, we present results on the overall GPK test score. Table 3 shows the means, standard errors of the means, and the standard deviations for the subsample of Future Primary Teachers pursuing a teaching career to teach grade 1 to 4 from LEK and TEDS-M. Whereas TEDS-M future primary

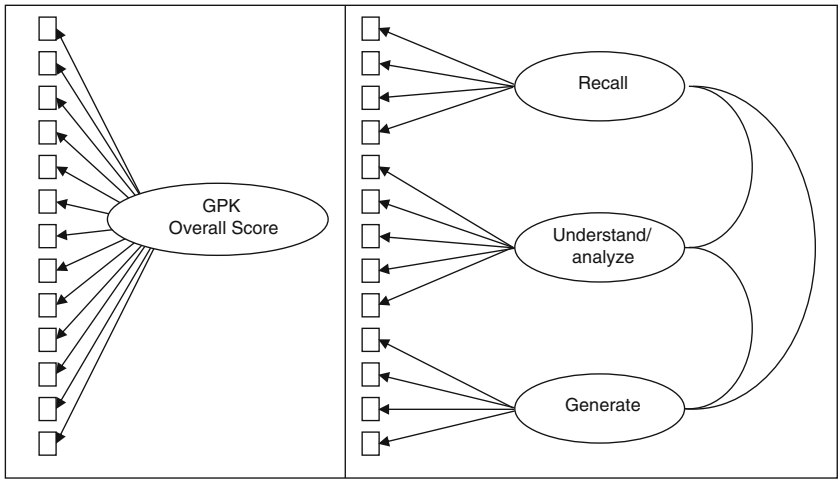

Figure 8. One-dimensional and three-dimensional modeling of general pedagogical knowledge (GPK)

teachers show a mean of 613 which is clearly above the international mean of 500 (for further details, see König & Blömeke, 2010b), future primary teachers' achievement with the same career path intentions who started in 2008 is more than two standard deviations lower. However, 2 years later in 2010 they show a much better test result which is only about one standard deviation lower than TEDS-M future teachers' achievement. All mean differences are statistically significant as the standard errors show, and they are all of high practical relevance.

Regarding future teachers' GPK measured by cognitive subscales as a second step of data analysis, obviously the pattern of the overall score is generally replicated (Figure 9) on all three subscales. Future primary teachers at the end of training outperform future teachers at the start of

TABLE 1

EAP reliability of general pedagogical knowledge overall score and cognitive subscales

	TEDS-M (König & Blömeke, 2010a, p. 272)	*LEK future primary teachers*
Overall score GPK	0.86	0.90
Cognitive subscales		
1. Recall	0.77	0.77
2. Understand/analyze	0.83	0.85
3. Generate	0.69	0.68

TEDS-M Teacher Education and Development Study in Mathematics, *LEK* Longitudinal Survey of Student Teachers' Pedagogical Competencies

TABLE 2

Latent intercorrelations of general pedagogical knowledge cognitive subscales

	TEDS-M (König & Blömeke, 2010a, p. 273)		*LEK future primary teachers*	
	1	*2*	*1*	*2*
1. Recall				
2. Understand/analyze	0.76		0.65	
3. Generate	0.69	0.83	0.44	0.56

TEDS-M Teacher Education and Development Study in Mathematics, *LEK* Longitudinal Survey of Student Teachers' Pedagogical Competencies

training, but after two years training future teachers show better test results though still performing lower than future teachers at the end of training. However, as Cohen's d as a measure of effect size (Table 4) indicates, the practical relevance of the mean differences on the subscales differs. Applying the classification according to which Cohen's $d \geq 0.2$ is of small, $d \geq 0.5$ is of medium, and $d \geq 0.8$ is of big effect size, mean differences on the subscale "Recall" can be interpreted as a continuous knowledge gain since mean differences are large between starting and having been trained for two years ($d = 1.21$) as well as between the latter stage and the end of training ($d = 1.44$). There are different results on the other two subscales. On the subscale "Understand/analyze," future teachers show a large knowledge gain over the first two years ($d = 1.33$) resulting in average achievement comparatively close to that of future teachers at the end of training ($d = 0.37$). By contrast, future teachers only show a medium effect size knowledge gain on the subscale "Generate" over the first two years of training ($d = 0.59$), but the mean difference between having been trained for two years and the end of training is very large again ($d = 1.29$).

TABLE 3

Overall general pedagogical knowledge test score

	M	*SE*	*SD*
Start of training (LEK)	372	7,6	130
After two years training (LEK)	525	7,1	98
End of training (TEDS-M)	613	5,3	84

TEDS-M Teacher Education and Development Study in Mathematics, *LEK* Longitudinal Survey of Student Teachers' Pedagogical Competencies

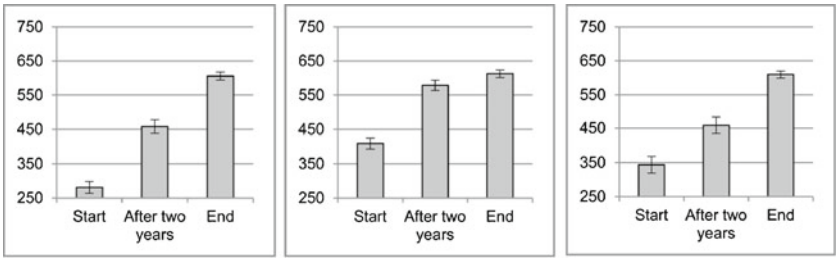

Figure 9. Means and 95 % confidence interval for future teachers' cognitive subscales scores

Results on the Level of Test Items

To illustrate the findings reported for the GPK scales of cognitive dimensions, it is worthwhile to provide information on the level of single test items that are characteristic for the two subscales "Understand/analyze" and "Generate."

Table 5 shows the percentage of future teachers from each state of training who answered scored items A to E of item example 1, a typical item measuring the cognitive process of "Understand/analyze." After two years of training, the proportion increase is about 30 % per item (with the exemption of item E). By contrast, the difference between the state of having been trained for two years and the end is only about 10 % per item.

Results for the second item example are shown in Table 6. When analyzing a lesson, presumably future teachers are increasingly relating to questions concerned with the process and the output of the lesson, such as teaching methods and student achievement, between having been trained for two years and the end of training, rather than between the start of training and having been trained for two years. Interestingly, it is the second criterion (input of the lesson) which is most strongly focused on as early as over the two years of training, whereas surprisingly the first criterion (context of the lesson) has very little meaning even for the future teachers at the end of training. [6]

TABLE 4

Cohen's *d* for mean differences in GPK cognitive subscales

	Recall	*Understand/analyze*	*Generate*
Start vs. after two years	1.21	1.33	0.59
Start vs. end	2.89	1.86	1.83
After two years vs. end	1.44	0.37	1.29

TABLE 5

Percentage (standard errors) of future teachers who answered scored items of item example 1 correctly

Test item	*Start*	*After two years*	*End*
A	50.3 (3.6)	84.4 (3.3)	93.7 (1.4)
B	46.2 (3.5)	75.4 (3.9)	85.2 (2.4)
C	52.8 (3.5)	89.3 (2.8)	96.8 (1.3)
D	44.2 (3.5)	71.3 (4.1)	80.3 (3.1)
E	23.6 (3.0)	32.0 (4.2)	39.5 (3.7)

Finally, Table 7 shows the percentage of future teachers from each state of training who answered scored items on item example 3, another item measuring the cognitive process of "generating." The proportion significantly increases between after two years of training and the end of training on the scored items (1) and (3), whereas no statistically significant difference can be found between the start and after two years of training, as the standard errors show. Again, it is interesting that teacher-centered methods are well known even to a substantial number of future teachers who have just started their teacher education, whereas student-centered methods providing students with feedback on their learning process are known by just one third of those who have finished their training.

Discussion

Teacher education systems worldwide are confronted with the essential question of how to provide appropriate opportunities for learning that foster

TABLE 6

Percentage (standard errors) of future teachers who answered scored items of item example 2 correctly

Four criteria addressed by scored items	*Start*	*After two years*	*End*
(1) Context of the lesson (e.g. prior knowledge of students)	16.0 (2.7)	22.2 (3.9)	23.5 (2.7)
(2) Input of the lesson (e.g. objectives of the lesson)	48.6 (3.7)	64.1 (4.5)	80.4 (2.7)
(3) Process of the lesson (e.g. teaching methods used)	70.7 (3.4)	72.6 (4.1)	79.3 (3.0)
(4) Output of the lesson (e.g. student achievement)	42.0 (3.7)	55.6 (4.6)	78.9 (3.1)

TABLE 7

Percentage (standard errors) of future teachers who answered scored items of item example 3 correctly

Three criteria addressed by scored items	*Start*	*After two years*	*End*
(1) Teacher-centered (e.g. oral inquiry by the teacher)	37.3 (3.6)	38.8 (4.5)	58.4 (2.9)
(2) Interactive (e.g. discussion of work results)	5.6 (1.7)	7.8 (2.5)	12.6 (1.9)
(3) Student-centered (e.g. self-evaluation by students)	8.5 (2.1)	13.8 (3.2)	32.4 (3.3)

both future teachers' theoretical and practical knowledge and that adequately enable future teachers to connect the theoretical and practical issues of teaching. In this article, data analysis was conducted to investigate how future teachers acquire GPK as a central component of teacher knowledge in the course of initial teacher education, exemplified by future teachers undergoing teacher education in Germany, where initial teacher education is divided into a first phase with a heavy focus on theoretical, academic study, and a second phase where intending future teachers learn how to perform and apply their theoretical knowledge in the classroom.

Findings show that the more advanced future teachers are in the course of their initial teacher education, the better they perform in the test measuring GPK. Mean differences between the three groups of primary future teachers (beginning, after two years, and end of teacher training) are statistically significant and practically relevant. This corresponds with the claim that teacher education in Germany is effective and the conceptualization of the test instrument is curricular valid with regard to German teacher education. Against this background, findings of TEDS-M international research on GPK (König et al., 2011) are strengthened since future teachers' GPK tested in TEDS-M most likely results from the learning processes in teacher education programs. Although the data analysis provided here is based on quasi-longitudinal modelling since a panel sample covering the whole length of teacher education from beginning to end with several occasions of measurement was not available, it is very improbable that TEDS-M future teachers had possessed the GPK as early as when they started their initial teacher education. However, findings presented here are limited insofar that we can interpret mean differences as a future teacher knowledge gain rather than drawing conclusions from the relevant statistical longitudinal modelling.

Since the German initial teacher education's first phase tends to foster future teachers' theoretical knowledge predominantly, we expected that future teachers acquired GPK related to the first two cognitive dimensions of the test instrument (recalling and understanding/analyzing) between the first occasion of measurement (entry to teacher education) and the second occasion of measurement after 2 years. By contrast, when comparing test results between future teachers having been educated for 2 years only and future teachers in their last year of training, we expected better test results particularly with regard to the third cognitive dimension (generating) to the benefit of the latter group. The findings show this is generally true. The typical features of first and second phase opportunities to learn were mirrored by the relative strengths and weaknesses of future teachers' GPK when groups at different stages were compared. Obviously, university teacher education is associated with the GPK growth related to understanding or analyzing a concept, a specific term or a phenomenon, whereas the practical second phase increasingly offers opportunities to learn where future teachers apply their declarative and conceptual GPK in order to generate concrete strategies on how they could solve a typical classroom situation problem. This clearly corresponds to the intentions of the two-phase teacher education system in Germany (Terhart, 2003; Messner, 2004) and the functions of its different institutions in universities (first phase), on the one hand, and *Studienseminare* of the federal states (second phase) on the other hand. What we did not expect was the ongoing growth of the first GPK cognitive dimension (recalling) which describes future teachers' retrieving information from long-term memory. Presumably during the second phase, future teachers are repeatedly confronted with this cognitive process, e.g. when planning a lesson, they are being given the opportunity to recall what they might have learned earlier during teacher education leading to the cognitive processes of consolidating their declarative GPK.

Moreover, these findings can be appropriately discussed with regard to the research on teacher expertise. As Berliner (2004, p. 205) points out using a heuristic model of teacher development, student teachers such as those sampled in the LEK-study (i.e. preservice teachers at the beginning of their teacher training) are considered to be novices. "At this stage, the commonplaces of an environment must be discriminated, the elements of the tasks to be performed need to be labeled and learned, and the novice must be given a set of context free rules." Applied to the findings here, the

declarative-conceptual knowledge gain measured between the beginning of teacher education and the time point when future teachers have been trained for 2 years widely corresponds with Berliner's description of what student teachers at the novice stage ordinarily have to master. Looking at the next stage, the "advanced beginner" (Berliner, 2004, p. 206) has already learned much of "the objective facts and features of the situations" and increasingly "experience is gained" that "can become melded with verbal knowledge, where episodic and case knowledge is built up." With regard to the situation of second phase pre-service teachers in Germany who regularly have to teach for 1.5 to 2 years, TEDS-M future teachers can plausibly be assigned to this stage. Again, applied to the findings here, TEDS-M future teachers' higher performance on the subscale "Generate" mirrors the qualitative difference of their knowledge distinguishing them from the "novice stage" and the "advanced beginner stage". To correctly respond to test items measuring the cognitive dimension of "generate," specific experience, that student teachers considered to be novices have not gained yet, is needed.

So we assume in-school opportunities to learn during teacher education, which enable future teachers to teach in school, are decisive when looking at the acquisition of the procedural elements of GPK.[7] In many teacher education systems worldwide, in-school opportunities to learn are to help future teachers to apply concepts, theories, etc. they have learned in the academic setting, and they also serve as a context where future teachers learn to teach (Wilson, Floden & Ferrini-Mundy, 2001) in order to act in a community of practice (Lave & Wenger, 1998). In Germany, these two functions are structurally related to each phase. For example, intentions of in-school opportunities to learn that are accompanied by university courses are frequently related to student teachers' academic reflection processes, whereas the intention to support future teachers' acquisition of classroom management routines or mastering of lesson and unit planning is typical of the curriculum of the *Studienseminare* of the second phase. In other countries with more integrative teacher education systems, student teachers might not experience structural differences as in Germany, but they also attend various forms of in-school opportunities to learn that underlie similar intentions to those we have outlined for the in-school opportunities to learn in Germany. Therefore, although we cannot generalize the results of our study to other countries' teacher education systems, implication of the results might be relevant for other systems as well.

What recommendations for future policy directions with respect to teacher education can be given? With regard to general pedagogy as a component of teacher education programs, broad claims about its uselessness as well as about what future teachers need to know at the end of their training have been made and linked with requests either to eliminate this component or to structure it in a new way (Grossman, 1992; Kagan, 1992). Although a quality evaluation of the sequence of theory before practice in initial teacher education is difficult, the findings presented here allow us to conclude that initial teacher education is effective. General intentions of the two phases find expression in the future teachers' GPK. Thus, the analysis carried out can be regarded as a starting point from which other initial teacher education systems, e.g. with a different linkage of theoretical and practical opportunities to learn (such as primary teacher education in Austria), can be evaluated by empirical testing of pre-service teachers in the course of their program in order to generate comparative information on how future teachers' knowledge acquisition might be fostered best in the future. Teacher knowledge research has just started to envisage such challenges to finally contribute to the development and improvement of a high quality of education and school system.

NOTES

[1] 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 analyses prepared for this paper and the views expressed are those of the authors and do not necessarily reflect the views of the IEA or the funding agencies.

[2] The LEK-study was funded by the German Research Foundation (DFG, KO 3947/3-1).

[3] Since we plan to use the test instrument in future studies, we do not include more than these item examples in this article. In case there is serious interest by other researchers doing research in the field of the topic under study, we will be pleased to provide them specifically with a complete documentation of a shorter version of the instrument including various materials such as about half of the test item pool, coding rubrics, scoring instruction, empirically based information on item parameters, and test booklets. This documentation (König & Blömeke, 2010c) allows other researchers to use a shorter version of the GPK test independently from the authors who developed it.

[4] Another psychometric indicator is the item fit statistics. The weighted mean squares mainly ranged from 0.80 to 1.20 with a few exceptions, which is a good result (Adams & Wu, 2002; Wright, Linacre, Gustafsson, Martin-Loff, 1994). Exceptions occurred when an item could not be excluded for conceptual reasons. Then, we accepted weighted mean squares ranging from 0.75 to 0.79 and from 1.21 to 1.25.

[5] This is a re-scaling since LEK test data were scaled before independently from TEDS-M data (König & Seifert, 2012). Findings show the test also serves as a reliable and valid measure when applied to pre-service teachers at a stage (beginning of initial teacher education and after 2 years of training) that is different to the stage of TEDS-M future teachers who were tested at the end of initial teacher education. So invariance of testing with regard to different cohorts of pre-service teachers is assured, a very important precondition for the analysis conducted in this article. This examination also included looking at anchor item parameter sets which were sufficiently invariant across the samples.

[6] As the standard errors show, differences of percentages of future teachers who answered the first scored item are not statistically significant.

[7] This assumption might be further confirmed by results from the LEK study. In a detailed longitudinal analysis (König, 2012, König & Seifert, 2012), the effects of a 2-week practicum during the first phase were investigated. The practicum did not require student teachers to teach in class. However, student teachers, who had used their practicum as an opportunity to teach for the first time, showed a substantial increase on the GPK subscale "generating" ($d = 0.55$), whereas their fellow future teachers who had attended the practicum as well, but who had not taught in class, did not show an increase on that subscale. However, there were no differences in the increase of the GPK overall score of the two groups. The effect was controlled for GPA and intrinsic motivations for teaching.

References

Adams, R. & Wu, M. (Eds.). (2002). *PISA 2000 technical report.* Paris, France: OECD. Retrieved from http://www.oecd.org/dataoecd/53/19/33688233.pdf. Accessed 24 April 2013.

Anderson, J. R. (1982). Acquisition of cognitive skills. *Psychological Review, 89*, 369–406.

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.

Baer, M., Dörr, G., Fraefel, U., Kocher, M., Küster, O., Larcher, S., . . .Wyss, C. (2007). Werden angehende Lehrpersonen durch das Studium kompetenter? [Do future teachers acquire competencies through studying?] *Unterrichtswissenschaft, 35*(1), 15–47.

Ball, D. L., Thames, M. H. & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59*(5), 389–407.

Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften [Keyword: Professional competence of teachers]. *Zeitschrift für Erziehungswissenschaft, 9*(4), 469–520.

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 Society, 24*(3), 200–212.

Bond, T. G. & Fox, C. M. (2007). *Applying the Rasch Model. Fundamental measurement in the human sciences* (2nd ed.). Mahwah, NJ: Erlbaum.

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. (1997). Kompetenzen, Funktionen und unterrichtliches Handeln des Lehrers [Skills, functions, and curricular activities of the teacher]. In F. E. Weinert (Ed.), *Enzyklopädie der Psychologie: Psychologie des Unterrichts und der Schule [Encyclopedia of psychology: Psychology of teaching and the school], vol. 3* (pp. 177–212). Göttingen, Germany: Hogrefe.

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, Netherlands: Elsevier.

Clift, R. T. & Brady, P. (2005). Research on methods courses and field experiences. In M. Cochran-Smith & K. M. Zeichner (Eds.), *Studying teacher education. The report of the AERA Panel on Research and Teacher Education* (pp. 309–424). Mahwah, NJ: Erlbaum.

Dann, H.-D. (2000). Lehrerkognitionen und Handlungsentscheidungen. [Teacher cognition and action decisions]. In M. K. W. Schweer (Ed.), *Psychologie der Lehrer-Schüler-Interaktion [Psychology of teacher–student-interaction]* (pp. 79–108). Opladen: Leske and Budrich.

Fenstermacher, G. D. (1994). The knower and the known: The nature of knowledge in research on teaching. *Review of Research in Education, 20*, 3–56.

Gerstenmaier, J. & Mandl, H. (2000). Wissensanwendung im Handlungskontext: Die Bedeutung intentionaler und funktionaler Perspektiven für den Zusammenhang von Wissen und Handeln. [Applications of knowledge in the context of performance: The meaning of intentional and functional perspectives for the correlation of knowledge and performance]. In H. Mandl & J. Gerstenmaier (Eds.), *Die Kluft zwischen Wissen und Handeln. Empirische und theoretische Lösungsansätze [The gap between knowledge and performance. Empirical and theoretical approaches]* (pp. 289–322). Göttingen: Hogrefe.

Good, T. L. & Brophy, J. E. (2007). *Looking in classrooms*. Boston, MA: Allyn & Bacon.

Grossman, P. L. (1992). Why models matter: An alternate view on professional growth in teaching. *Review of Educational Research, 62*(2), 171–179.

Grossman, P. L. & Richert, A. E. (1988). Unacknowledged knowledge growth: A re-examination of the effects of teacher education. *Teaching and Teacher Education, 4*(1), 53–62.

Gruber, H. & Rehrl, M. (2005). *Praktikum statt Theorie? Eine Analyse relevanten Wissens zum Aufbau pädagogischer Handlungskompetenz.* [Practicum instead of theory? Analysis relevant knowledge for the acquisition of pedagogical performance] Forschungsbericht 15 des Instituts für Pädagogik, Universität Regensburg.

Hascher, T. (2006). Veränderungen im Praktikum–Veränderungen durch das Praktikum. Eine empirische Untersuchung zur Wirkung von schulpraktischen Studien in der Lehrerbildung [Change in practicum–change by practicum. Empirical investigation on the effectiveness of in-school opportunities to learn in teacher training]. *Zeitschrift für Pädagogik, 51*(Beiheft), 130–148.

Hatano, G. & Inagaki, K. (1986). Two courses of expertise. In H. W. Stevenson, H. Azuma & K. Hakuta (Eds.), *Child development and education in Japan: A series of books in psychology* (pp. 262–272). New York: Freeman.

Helmke, A. (2003). *Unterrichtsqualität erfassen, bewerten, verbessern [Understand, evaluate, and improve teacher quality]*. Seelze, Germany: Kallmeyer.

Jonsson, A. & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review, 2*(2), 130–144.

Kagan, D. M. (1992). Professional growth among preservice and beginning teachers. *Review of Educational Research, 62*(2), 129–169.

Klafki, W. (1985). *Neue Studien zur Bildungstheorie und Didaktik. Beiträge zur kritisch-konstruktiven Didaktik [New studies to education theory and didactics: Contributions to the critical constructional didactics]*. Weinheim, Germany: Beltz.

Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., Reiss, K., Riquarts, K., Rost, J., Tenorth, H.-E. & Vollmer, H. J. (Eds.). (2003). *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise. [On the development of national standards. An expertise]. Bildungsforschung. Bd. 1*. Bonn: Verlag.

Kolbe, F.-U. & Combe, A. (2004). Lehrerbildung [Teacher education]. In W. Helsper & J. Böhme (Eds.), *Handbuch der Schulforschung [Handbook of school research]* (pp. 853–877). Wiesbaden: VS.

König, J., & Blömeke, S. (2009). Pädagogisches Wissen von angehenden Lehrkräften: Erfassung und Struktur von Ergebnissen der fachübergreifenden Lehrerausbildung [Pedagogic knowledge of prospective teachers: Acquisition and structure of the results of cross-disciplinary teacher education]. *Zeitschrift für Erziehungswissenschaft, 12*(3), 499–527.

König, J. (2010). Lehrerprofessionalität – Konzepte und Ergebnisse der internationalen und deutschen Forschung am Beispiel fachübergreifender, pädagogischer Kompetenzen. [Teacher Professionalism - Concepts and Findings of International and German Research exemplified with Pedagogical Competencies.] In J. König & B. Hofmann (Eds.), *Professionalität von Lehrkräften – Was sollen Lehrkräfte im Lese- und Schreibunterricht wissen und können?* (pp. 40–105). [Teacher Professionalism - What Should Teachers should Know and Be Able to for Teaching Reading and Writing?] Berlin: DGLS.

König, J., & Blömeke, S. (2010a). Messung des pädagogischen Wissens: Theoretischer Rahmen und Teststruktur [Measurement of educational knowledge: Theoretical framework and test structure]. In S. Blömeke, G. Kaiser, & R. Lehmann (Eds.), *TEDS-M 2008 – Professionelle Kompetenz und Lerngelegenheiten angehender Primarstufenlehrkräfte im internationalen Vergleich* [Professional competence and learning opportunities of prospective primary school teachers in an international comparison] (pp. 253–273). Münster, Germany: Waxmann.

König, J., & Blömeke, S. (2010b). Pädagogisches Wissen angehender Primarstufenlehrkräfte im internationalen Vergleich [Pedagogic knowledge of prospective primary school teachers in an international comparison]. In S. Blömeke, G. Kaiser, & R. Lehmann (Eds.), *TEDS-M 2008 – Professionelle Kompetenz und Lerngelegenheiten angehender Primarstufenlehrkräfte im internationalen Vergleich* [Professional competence and learning opportunities of prospective primary school teachers in an international comparison] (pp. 275–296). Münster, Germany: Waxmann.

König, J., & Blömeke, S. (2010c). Pädagogisches Unterrichtswissen (PUW) [General Pedagogical Knowledge for Teaching]. In *Dokumentation der Kurzfassung des TEDS-M-Testinstruments zur Kompetenzmessung in der ersten Phase der Lehrerausbildung* [Documentation of the short version of the TEDS-M-Test instrument for competence assessment in the first phase of teacher training]. Berlin, Germany: Humboldt University of Berlin.

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. (2012). Zum Einfluss der Schulpraxis im Lehramtsstudium auf den Erwerb von pädagogischem Wissen: Spielen erste Unterrichtsversuche eine Rolle? [On the Influence of School Practice in Teacher Education Effecting the Acquisition of Pedgogical Knowledge: Does Initial Teaching Matter?] In T. Hascher & G. H. Neuweg (Eds.), *Forschung zur (Wirksamkeit der) LehrerInnenbildung* (pp. 143-159). [Research on (the Effectiveness of) Teacher Education. Wien: LIT-Verlag.

König, J., & Seifert, A. (Eds.). (2012). Lehramtsstudierende erwerben pädagogisches Professionswissen. Ergebnisse der Längsschnittstudie LEK zur Wirksamkeit der erziehungswissenschaftlichen Lehrerausbildung. [Student Teachers Acquire Professional Pedagogical Knowledge. Findings from Longitudinal Study LEK on the Effectiveness of General Pedagogy in Teacher Education.] Münster: Waxmann.

König, J., Kaiser, G., & Felbrich, A. (2012). Zum Zusammenhang von Wissen und Überzeugungen am Ende der Lehrerausbildung: Spiegelt sich pädagogisches Wissen in den Kompetenzselbsteinschätzungen angehender Lehrkräfte? [Is Pedagogical Knowledge Reflected in the Competence-Related Self-Assessments of Future Teachers? On the inter-relation between knowledge and beliefs after completing teacher education.] *Zeitschrift für Pädagogik, 58*(4), 476–491.

König, J. & Blömeke, S. (in press). 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 the Seventeen TEDS-M Countries.

Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (Eds.). (2011). *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV [Professional competence of teachers: Finding from COACTIV]*. Münster: Waxmann.

Larcher, S. & Oelkers, J. (2004). Deutsche Lehrerbildung im internationalen Vergleich [German teacher education: An international comparison]. In S. Blömeke, P. Reinhold, G. Tulodziecki & J. Wildt (Eds.), *Handbuch Lehrerausbildung [Handbook of teacher training]* (pp. 128–150). Bad Heilbrunn, Germany: Klinkhardt.

Lave, J. & Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge: Cambridge University Press.

Lortie, D. (1975). *Schoolteacher: A sociological study*. Chicago: The University of Chicago Press.

McDonald, J. P. (1992). *Teaching: Making sense of an uncertain craft*. New York, NY: Teachers College Press.

Messner, R. (2004). Leitlinien einer phasenübergreifenden Lehrerbildung [Guidelines of phases-overarching teacher education]. *Seminar Themenheft Lehrerbildung und Schule, 4*, 9–27.

Oser, F. & Oelkers, J. (Eds.). (2001). *Die Wirksamkeit der Lehrerbildungssysteme [The effectiveness of teacher education systems]*. Chur, Switzerland: Ruegger.

Putnam, R. T. & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? *Educational Researcher, 29*(1), 4–15.

Schmidt, W. H., Cogan, L. & Houang, R. (2011). The role of opportunity to learn in teacher preparation: An international context. *Journal of Teacher Education, 62*(2), 138–153.

Schmidt, W. H., Tatto, M. T., Bankov, K., Blömeke, S., Cedillo, T., Cogan, L., Han, S.-I., Houang, R., Hsieh, F.-J., Paine, L., Santillan, M. N. & Schwille, J. (2007). *The preparation gap: Teacher education for middle school mathematics in six countries—mathematics teaching in the 21st century (MT21)*. East Lansing. Retrieved December 12, 2007, from http://usteds.msu.edu/related_research.asp.

Schulte, K. (2008). *Selbstwirksamkeitserwartungen in der Lehrerbildung–Zur Struktur und dem Zusammenhang von Lehrer-Selbstwirksamkeitserwartungen, Pädagogischem Professionswissen und Persönlichkeitseigenschaften bei Lehramtsstudierenden und Lehrkräften* [Self-efficacy in teacher education—on the structure and the relation of teacher self-efficacy, pedagogical knowledge, and personality traits of pre-service and in-service teachers]. University of Göttingen. Retrieved from: http://webdoc.sub.gwdg.de/diss/2008/schulte/schulte.pdf.

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.

Slavin, R. E. (1994). Quality, appropriateness, incentive, and time: A model of instructional effectiveness. *International Journal of Educational Research, 21*, 141–157.

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., Schwille, J., Senk, S., Ingvarson, L., Rowley, G., Peck, R., Bankov, K., Rodriguez, M. & Reckase, M. (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)*. Retrieved April 10, 2012, from http://www.iea.nl/fileadmin/user_upload/Publications/Electronic_versions/IEA_TEDS-M.pdf.

Terhart, E. (1993). Pädagogisches Wissen. Überlegungen zu seiner Vielfalt, Funktion und sprachlichen Form am Beispiel des Lehrerwissens [Pedagogical knowledge. On its diversity, function, and linguistic form exemplified by teacher knowledge]. In J. Oelkers & H.-E. Tenorth (Eds.), *Pädagogisches Wissen [Pedagogical knowledge]* (pp. 129–141). Weinheim/Basel: Beltz.

Terhart, E. (2003). Lehrerbildung nach PISA. Welche Konsequenzen kann man aus den aktuellen Leistungsvergleichsstudien für die Lehrerbildung ziehen? [Teacher education after PISA. Which consequences can one draw from current comparative assessments for teacher education?]. In H. Merkens (Ed.), *Lehrerbildung in der Diskussion [Discussing teacher education]. Schriften der Deutschen Gesellschaft für Erziehungswissenschaften* (pp. 167–177). Opladen: Leske + Budrich.

Tulodziecki, G., Herzig, B. & Blömeke, S. (2004). *Gestaltung von Unterricht. Eine Einführung in die Didaktik [Design of instruction: An introduction to teaching]*. Bad Heilbrunn, Germany: Klinkhardt.

von Davier, A. A., Carstensen, C. H. & von Davier, M. (2006). Linking competencies in horizontal, vertical, and longitudinal settings and measuring growth. In J. Hartig, E. Klieme & D. Leutner (Eds.), *Assessment of competencies in educational contexts* (pp. 53–80). Göttingen: Hogrefe.

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.

Wilson, S. M. & Berne, J. (1999). Teacher learning and the acquisition of professional knowledge: An examination of research on contemporary professional development. *Review of Research in Education, 24*, 173–210.

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

Wright, B. D., Linacre, M., Gustafsson, J.-E. & Martin-Loff, P. (1994). Reasonable mean square fit values. *Rasch Measurement Transactions, 8*, 370.

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

Empirical School Research
University of Cologne
Gronewaldstr. 2, 50931 Cologne, Germany
E-mail: johannes.koenig@uni-koeln.de