

Running head: PROCESS MEDIATES STRUCTURE

Process mediates structure: Relation of preschool teacher education to preschool teachers' knowledge

Sigrid Blömeke (corresponding author)

Centre for Educational Measurement at the University of Oslo (CEMO)

N – 0318 Oslo, Norway

Email: sigribl@cemo.uio.no, Tel. ++47 464 18 755

Lars Jenßen, Marianne Grassmann

Humboldt University of Berlin

D – 10099 Berlin, Germany

Simone Dunekacke

Leibniz Institute for Science and Mathematics Education

D – 24118 Kiel, Germany

Hartmut Wedekind

Alice Salomon University of Applied Science

D – 12627 Berlin, Germany

ACKNOWLEDGEMENT

The research project „Structure, level and development of prospective preschool teachers‘ competencies in mathematics” (KomMa) was funded by the German Federal Ministry of Education and Research (BMBF) as part of the research initiative “Modelling and measuring competencies in higher education (KoKoHs)” (FKZ 01PK11002A).

Submission date: November 18, 2015

Revised version: March 18, 2016

Second revision: June 20, 2016

Final version with author identifiers: June 27, 2016

ABSTRACT

Data about processes and outcomes of preschool teacher education is scarce. This paper examines the opportunities to learn (OTL) of prospective preschool teachers (N=1,851) at different types and stages of preschool teacher education and their relation to general pedagogical knowledge (GPK), mathematics pedagogical content knowledge (MPCK) and mathematical content knowledge (MCK) with standardized tests. Process indicators in terms of OTL and outcome indicators in terms of knowledge varied substantially across teacher education types and stages. Controlling for preschool teachers’ background, multi-level models revealed that OTL in general pedagogy and mathematics pedagogy provided during teacher education were significantly related to GPK and MPCK. Effect sizes reached up to two thirds of a standard deviation. OTL were in turn significantly related to the type of institution that offered a program in favor of pedagogical colleges compared to vocational schools. OTL were also significantly

related to program stage in favor of the last year of preschool teacher education compared to the beginning. Process characteristics in terms of OTL mediated fully or partly structural characteristics of teacher education such as type of institution or program stage. These results suggest that the OTL provided are more important than whether prospective preschool teachers were at the beginning or the end of their program or whether they were prepared at vocational schools or pedagogical colleges (although entrance differences have still be taken into account). It may be an important responsibility of policy makers then to ensure that all prospective preschool teachers receive sufficient OTL.

Keywords: preschool teachers, teacher education, early childhood education, pedagogical knowledge, content knowledge, educational effectiveness, opportunity to learn, multi-level modelling

Educational Impact and Implications Statement

This article shows that opportunities to learn general pedagogy and mathematics pedagogy by prospective preschool teachers during their teacher education program were related strongly related to their general pedagogical and mathematics pedagogical content knowledge. How many opportunities to learn prospective teacher received was in turn related to the type and stage of a teacher education program in favor of pedagogical colleges when compared to vocational schools and in favor of the last year of preschool teacher education when compared to the beginning of their education. These findings provide important information to understanding how preschool teachers gain their professional knowledge, and these results can assist policy makers in deciding about how to improve preschool teacher education. The results suggest that opportunities to learn provided during preschool teacher education may be more important for

knowledge acquisition than more distal factors such as the type of institution where prospective teachers are prepared. This may in turn suggests that it may be worthwhile to focus reforms of preschool teacher education more directly on opportunities to learn instead of on less direct structural changes.

Highlights

- Standardized tests of prospective preschool teachers' knowledge were developed.
- Objectivity, reliability and content, construct and criterion validity was confirmed.
- Domain-specific opportunities to learn were strongly related to GPK and MPCK.
- Type of teacher education institution and program stage were related to OTL.
- Relation of teacher education structure was partly mediated through processes.

Research on effects of preschool teacher education on prospective preschool teachers' knowledge and skills has so far mostly been restricted to distal indicators of teacher knowledge such as degrees or licensing (Whitebook, Gomby, Bellm, Sakai, & Kipris, 2009). The same applies to studies examining the opportunities to learn (OTL) of prospective preschool teachers, in the present paper defined as content coverage providing the chance to gain the knowledge and skills necessary to succeed with fostering the development of preschool-age children (i.e., 3 to 6 years of age). In most studies, OTL were operationalized through distal indicators such as the length or the type of a preschool teacher education program (Bogard, Traylor, & Takanishi, 2008). Corresponding to the state of research on primary and secondary school teacher education (Abell Foundation, 1992; Darling-Hammond, 2001), results from research on preschool teacher education have been contradictory. Whereas some studies have established significant relations between preschool teacher education and preschool teachers' knowledge and skills or long-term outcomes such as children's development (Burchinal et al. 2002; Howes et al., 1992; Tout et al., 2005; Whitebook et al., 2009), other studies have failed to establish relations (Early et al., 2007).

Most authors agree that this unsatisfactory state of research is due to problems with the measures that have been used. Degrees and licenses but also program length and types are rather imprecise (i.e., unreliable) indicators of the knowledge and skills that preschool teachers gain during their education or the OTL they encounter during teacher education. The meanings of the distal indicators depend on the specific norms and practices applied in different teacher education institutions (Carroll, 1963). Educational effectiveness research has revealed that the content that is covered (Berliner, 1985) and the time allocated to such OTL (Carroll, 1963) are at the core of teaching and learning (Travers & Westbury, 1989). Standardized instruments for

assessing preschool teacher educations' OTL and their outcomes in such a specific way are missing but urgently needed (Bogard et al., 2008; Early et al., 2007).

Furthermore, there is no systematic framework that is able to define the structure of preschool teachers' knowledge and skills and conceptualize their dimensions in more detail. Nor is there a systematic framework beyond institution-specific curricula that conceptualizes the OTL offered during preschool teacher education. Such frameworks are therefore urgently needed as well. The distal indicators currently in use (e.g., degree or program length) are only rough approximations. They are not indicative of specific *domains* such as reading or mathematics, let alone sufficiently specific with respect to details within these domains.

State of Research

A summary of the state of research on the relation between preschool teacher education and teacher knowledge reveals substantial holes. Furthermore, due to the lack in a shared understanding of the construct "OTL" and the lack of standardized measures different authors operationalize OTL differently which leads to some ambiguity in the following review as well.

A nationally representative U.S. study found that domain-specific OTL in mathematics, reading, or science are scarce during preschool teacher education because even at institutions of higher education, most programs focus on general pedagogical OTL (Early & Winton, 2001; see also Isenberg, 2000). Linguistic and cultural diversity or the education of children with disabilities were additional blind spots (Lobman, Ryan, & McLaughlin, 2005). In an analysis of preschool teachers' self-reports, another U.S. study correspondingly found that they did not feel sufficiently prepared to teach children with diverse backgrounds (Ryan, Ackerman, & Song, 2004).

If one takes into account research on professional development (PD) after initial teacher education, a clearer picture emerges. Hamre et al. (2012) found substantial effect sizes in the relation between OTL offered to preschool teachers and their ability to perceive the classroom accurately as assessed with a standardized video test as well as between OTL and the teachers' ability to support children's literacy skills as assessed with a standardized test. Similar results were found by Pianta et al. (2014). However, Piasta, Logan, Pelatti, Capps, and Petrill (2015) were able to provide evidence for PD effects on performance in preschools only in the domain of children's science learning but not in mathematics learning.

Thus, we have initial evidence that OTL matter in that they are related to preschool teachers' knowledge and skills. Despite frequent pleas for more research on the specific effects of preschool teacher education on teacher characteristics with standardized and domain-specific measures (Early et al., 2007; Whitebook et al., 2009), not many studies have undertaken this effort though. Whereas recently a large number of studies using direct, standardized, and domain-specific teacher assessments has been published on primary and secondary teacher education—confirming strong links between OTL during teacher education and teacher education outcomes in terms of prospective teachers' knowledge, which in turn predicted teaching performance and student achievement (Blömeke, Suhl, Kaiser & Döhrmann, 2012; Tatto et al., 2012; Voss, Kunter, & Baumert, 2015)—preschool teacher education is still a “black box.”

One particular blank spot exists with respect to the effectiveness of preschool teacher education below the tertiary level which applies to many developing countries but also to a range of Southern and Western European countries (Wallet, 2006). Such programs do not take place at institutions of higher education but at post-secondary or secondary vocational school.

Completion of high school is thus not necessarily a requirement for entering a preschool teacher education program.

In many countries, policy efforts have been undertaken to move preschool teacher education up to the tertiary level. A prominent example of this is the Head Start program in the US, which is directed toward providing high-quality preschools to low-income children. The program receives funding only on the condition that half of its preschool teachers hold a Bachelor's degree (Bassok, 2012). Graduates from Bachelor programs have thus become the main target population of preschool research (see, e.g., Early et al., 2007). However, with a few exceptions such as some Scandinavian countries and some states in the US, preschool teachers with a degree below the tertiary level are still the majority in the US (Bogard et al., 2008) and in many other countries, including Germany, which is the context of the present study.

To overcome the research gaps described above, the objective of this paper is to unpack the black box of “preschool teacher education” by examining the relation between domain-specific OTL provided during preschool teacher education and domain-specific outcomes in terms of preschool teachers' knowledge while controlling for their background characteristics. All knowledge dimensions were assessed in a standardized way in a multicohort, multigroup design to be able to include prospective preschool teachers from different types of teacher education institutions. The instruments were developed on the basis of a conceptual framework derived from educational effectiveness research, which will be presented in this paper as well. We paid particular attention to differences between institutions of higher education that award a Bachelor's degree to preschool teachers and vocational schools part of the secondary school level.

How important it is to clarify the relation between preschool teacher education and teacher knowledge is demonstrated in studies that focused on the relation between this knowledge and the cognitive development of children. Early et al. (2006) found that preschool teachers with a Bachelor's degree delivered higher mathematics-related instructional quality as indicated by standardized on-site observations and achieved stronger outcomes in a direct assessment of children's mathematical literacy than preschool teachers without such a degree. Preschool teachers' knowledge in mathematics, assessed directly with a standardized test, also significantly predicted their ability to perceive preschool situations appropriately and to perform instructional activities that support the development of children's mathematics literacy as assessed with a standardized video test (Dunekacke, Jenßen & Blömeke, 2015a). Evidence exists in other domains (e.g., reading literacy) as well (Connor, Morrison, & Slominski, 2006; Landry, Anthony, Swank, & Monseque-Bailey, 2009).

Preschool Teacher Education in the Context of Germany

Preschool education in Germany is voluntary and can be subdivided into institutions covering 1-to 3-year-olds and institutions covering 3-to-6-year-olds. Teachers of the latter represent the target population of this study. At this age, more than 90% of the children are enrolled at least part-time—mostly in morning sessions—although parents have to pay a small fee (Statistisches Bundesamt, 2014). Preschools are typically run by local municipalities, churches (mostly Protestant or Catholic), or charity organizations, and some are organized privately with a special pedagogical profile. Preschools are not part of the school system but of the child and youth welfare system. They are therefore assigned to ministries of family affairs instead of ministries of education in the 16 German states so that there is more emphasis on care than on formal education.

Play-based activities represent the norm for teacher-child interactions (Liegler, 2008). Preschools organize these activities either in fixed groups with one full-time preschool teacher (or equivalent part-time employees) assigned to about 10 children or in larger groups of variable sizes looked after by teams of preschool teachers. Because more and more evidence points to the relevance of child development prior to schooling for later student achievement (see, e.g., Duncan et al., 2006), the belief that it is important to foster young children's cognitive development has increased in recent years—in particular with respect to 3-to-6-year-olds. All 16 German states have recently implemented standards for preschools that present ambitious cognitive objectives with respect to early reading, mathematics, and science literacy. This means that teachers have to use the informal context of preschool more often than before to foster these abilities. However, a systematic accountability system to support the achievement of these ambitious objectives does not yet exist.

Due to society's increased awareness of the relevance of preschool education, parents have recently earned the right to send their children to preschool beginning at age 1 when the paid parenthood leave ends. If a municipality is not able to offer such a child a spot in a preschool, parents are reimbursed for the private daycare costs that exceed the small fee they would have to pay for a spot in a regular preschool.

Preschool teachers are trained differently in the 16 German states. Typically, a two-tiered system exists. The majority of preschool teachers (more than 90%) are trained at vocational schools that provide teacher education on the secondary or postsecondary level. This means that the entrance requirement is not completion of high school but of 9 or 10 years of general schooling followed by 2 to 4 years of vocational training in a care profession (or a similar type of education). In parallel, there are also pedagogical colleges that are part of the higher education

system and award a Bachelor's degree. 56 colleges existed at the time of our study in 2015. Students must have completed high school followed by a 6-to-12-month pedagogical internship to enter these colleges. Currently, only about 5% of preschool teachers have undergone this type of education, and the numbers are growing only slowly (Statistisches Bundesamt, 2014). The 16 German states are responsible for the preschool teacher education curricula; the 56 pedagogical colleges have the academic freedom to design their curricula so that the training conditions vary substantially across Germany.

Conceptual Framework

To the best of our knowledge, there is no conceptual framework that specifically describes the structure of preschool teachers' knowledge. To avoid a purely operational definition, we therefore applied basic educational-psychological dimensions of primary teachers' knowledge to preschool teachers but operationalized these on the basis of research on 3-to-6-year-old children's development and learning. This approach ensured connectivity between subsequent educational stages (Anders, 2012) so that we could examine the specifics of each one.

Preschool Teachers' Knowledge

According to Shulman (1986) and Weinert (2001), teacher knowledge is a multidimensional construct that includes general pedagogical knowledge, pedagogical content knowledge, and content knowledge. With respect to content, the present study was restricted to the domain of mathematics learning. Preschool teachers' knowledge then includes mathematics content knowledge (MCK), pedagogical content knowledge of how to foster mathematics learning in children between the ages of 3 and 6 (MPCK), and general pedagogical knowledge of how to organize the informal learning environment of preschool in general (GPK).

To define these dimensions in more detail, we conducted two systematic analyses of all preschool teacher education curricula from the 56 pedagogical colleges and the 16 federal states (for the vocational schools) as well as of all preschool standards implemented in the 16 federal states (Jenßen, Dunekacke & Blömeke, 2015). Preschool standards set by the 16 German states were used to describe the objectives of preschool with respect to children's mathematical learning. Preschool teacher education curricula were used to describe the OTL in mathematics, mathematics pedagogy, and general pedagogy offered to prospective preschool teachers at the different institutions in the 16 states. Construct maps (Wilson, 2005) summarized the results of the systematic analyses of preschool teacher education curricula and standards in terms of subdimensions and specific descriptors. These were used to represent the range of pedagogical and mathematical OTL and preschool objectives. During test development, these descriptors were operationalized with items that were represented in the majority of standards and curricula and were also supported by the literature (for detailed results, see the Appendix).

GPK includes general foundations from educational theory, psychology, and instructional research related to early childhood and learning processes of 3-to-6-year-olds (Blömeke, Jenßen, Dunekacke, Suhl, Grassmann & Wedekind, 2015). An OECD (2004) review of early childhood curricula in five countries revealed that the present framework is in alignment with discussions elsewhere. *MPCK* includes diagnosing children's developmental state in mathematics and designing an informal learning environment that fosters the mathematical learning of children between the ages of 3 and 6 (Dunekacke, Jenßen & Blömeke, 2015b). Again, this framework resembles discussions in other countries (NAEYC, 2009). *MCK* includes numbers, sets, and operations; shape, space, and change; quantity, measurement, and relations; data, combinatorics, and chance (Dunekacke, Jenßen & Blömeke, 2015a). Although developed in the national context

of Germany, this framework also reflects discussions that are taking place elsewhere (Clements, Sarama, & DiBiase, 2004; National Research Council, 2009).

To ensure that the tests also included different cognitive processes, a second framework was developed on the basis of cognitive psychology (Anderson & Krathwohl, 2001). On the one hand, the items had to assess the recalling, understanding, and applying of knowledge as well as the knowledge-based generation of strategies. On the other hand, they had to capture cognitive complexity in terms of the different numbers of cognitive steps necessary to solve an item as well as different types of problem representations (Embretson & Daniel, 2008). The two frameworks, the alignment of frameworks and measures, as well as the inferences to be drawn from these measures have been validated in a range of studies (Blömeke, Jenßen, Dunekacke, Suhl, Grassmann & Wedekind, 2015; Dunekacke, Jenßen & Blömeke, 2015a, b; 2015d, Jenßen, Dunekacke, Eid & Blömeke, 2015).

OTL Provided During Preschool Teacher Education

Characteristics of preschool teacher education that potentially have an effect on prospective preschool teachers' GPK, MPCK and MCK because of the differences in OTL provided, are the *type* of institution where a program takes place (in the present study: pedagogical college vs. vocational school) and the program *stage* (the beginning vs. end of a program). In samples of prospective primary and secondary teachers, there is evidence that these aspects of German teacher education matter in favor of longer programs—typically also requiring stronger entrance characteristics—on the one hand, and in favor of students at the end compared to students at the beginning of teacher education (Blömeke, Kaiser & Lehmann, 2008; Kleickmann et al., 2013). The sample of the present study will therefore be drawn according to

these structural characteristics of preschool teacher education, and we will test corresponding hypotheses (see below H2a, b).

However, such structural characteristics of institutions are proxies rather than direct measures of the teaching and learning processes going on. According to educational effectiveness research, OTL in terms of the content that was covered has to be taken into account (Berliner, 1985; Carroll, 1963). OTL reflect preschool teachers' chances to acquire GPK, MPCK and MCK. OTL item development followed the same conceptual framework as applied for the three knowledge tests (see Appendix).

With respect to primary and secondary teacher education, evidence exists that such domain-specific proximal measures of teaching and learning processes are significantly related to outcomes (Blömeke, Suhl & Kaiser, 2011; König, Blömeke, Paine, Schmidt & Hsieh, 2011; Blömeke, Suhl, Kaiser & Döhrmann, 2012). GPK was significantly related to OTL in general pedagogy, whereas OTL in mathematics were significantly related to MCK and MPCK. OTL in mathematics pedagogy were significantly related to MPCK only when MCK was not included. These results applied both to primary and to secondary teachers. The aim of the present study is to expand this state of research to prospective preschool teachers by testing corresponding hypotheses (see below H1a, b, c, d).

Results from educational effectiveness research also revealed that in addition to examining such direct OTL effects on outcomes, indirect effects also need to be examined—for example, whether distal predictors such as structural characteristics of teacher education are mediated by proximal process indicators (see with respect to primary teachers Scheerens & Blömeke, 2016). Such a hypothesis is applied to preschool teacher education in this study as well (see below H3).

Hypotheses

H1: We hypothesized significant positive relations between different domain-specific process indicators of teaching and learning during preschool teacher education and corresponding teacher education outcomes. More specifically, we hypothesized that OTL in general pedagogy would have a stronger positive relation to GPK than OTL in mathematics pedagogy or in mathematics would (H1a). At the same time, we hypothesized that OTL in mathematics pedagogy would have a stronger positive relation to MPCK than OTL in general pedagogy or in mathematics would (H1b). Finally, we hypothesized that OTL in mathematics would have a stronger positive relation to MCK than OTL in general pedagogy or in mathematics pedagogy would (H1c).

Furthermore, we hypothesized that OTL would predict knowledge in a similar way in all subpopulations, which means technically that the relations would be invariant across prospective preschool teachers at pedagogical colleges and vocational school as well as across students at the beginning and at the end of teacher education (H1d).

H2: We hypothesized that structural characteristics of preschool teacher education would significantly positively predict prospective preschool teachers' knowledge. More precisely, prospective preschool teachers from pedagogical colleges were hypothesized to have significantly higher MCK, MPCK, and GPK compared with students from vocational schools (H2a). In addition, we hypothesized that MCK, MPCK, and GPK would be higher at the end compared with the beginning of preschool teacher education in both institutions, indicating progress (H2b).

H3: Finally, we tested a mediation model (see Figure 1). We hypothesized that the relations of structural teacher education characteristics—type of institution and program stage—to

MCK, MPCK, and GPK would be at least partly if not fully mediated through process characteristics in terms of the respective domain-specific OTL.

Insert Figure 1 here

Method

Participants

The sample included 1,851 prospective preschool teachers from 86 classes in 44 teacher education institutions. Each class had between 6 and 82 students ($M = 21$). The 44 institutions included 31 of 516 vocational schools in Germany with a total of 67 classes ($M_{Stud/Class} = 20$, $SD = 7.9$, Range = 6 to 46) and 13 of 80 German pedagogical colleges with a total of 19 classes ($M_{Stud/Class} = 25$, $SD = 18.8$, Range = 6 to 82). From most institutions one class participated in the study but from a few vocational schools up to four classes participated.

The sample was drawn via personal contacts as a first step and by randomly contacting vocational schools and pedagogical colleges as a second step. Because preschool teacher education is the responsibility of the states in Germany, in this second step, care was taken to include all 16 states and to represent the larger states by including a larger number of schools from them than from the smaller states. Only a few institutions that we contacted were not willing to participate. Any institution that declined was replaced by another randomly drawn institution from the same state.

Because the relations of OTLs to outcomes on the one hand and the differences between preschool teacher education at the secondary and higher education levels on the other hand were important research foci, we included four groups that were tested at the same time (see Table 1): prospective preschool teachers at the end and at the beginning of teacher education at institutions offering secondary education (vocational schools) and higher education (pedagogical colleges).

Prospective preschool teachers in higher education were purposefully oversampled because otherwise the group would have been too small for scaling purposes given that it is a small minority of all prospective preschool teachers (about 10% only; Statistisches Bundesamt, 2014).

The consent of test takers was obtained by pointing out before the assessment started that participation in the study was voluntary and that beginning to fill out the forms was taken as consent. Those who did not want to participate were given the opportunity to leave the room at that moment. The instructions included the additional information that every participant could leave the room at any time and that consent could be withdrawn at any time until the tests were collected. None of the participants used this option, but it is possible that a small number of students did not come to school on the day of testing because they may have heard about it beforehand. Table 2 provides an overview of the sample's major characteristics.

Insert Tables 1 and 2 here

The descriptive statistics were in line with our expectations and the demographics of the target population. The teachers in our sample who were at the end of preschool teacher education were 2 (vocational schools) or 4 (pedagogical colleges) years older than those who were at the beginning. Female teachers represented the majority in all four subgroups, and teachers' language background was almost always German. The biggest differences existed with respect to the two indicators of prior knowledge (school degree and number of years of mathematics in school) and the two indicators of socioeconomic background (mother's education and number of books at home). On each of the four indicators, the participants from vocational schools were at a disadvantage compared with the higher education students. The latter group reflects the German average with respect to mother's education (Statistisches Bundesamt, 2010, p. 26).

Measures

Preschool teacher education outcomes: GPK, MPCK, and MCK. Test development was applied according to the conceptual framework described (for details, see the Appendix). A large item pool ($n = 117$) was developed in a joint effort between academic and practical experts from preschool mathematics, mathematics pedagogy, and general pedagogy. Item selection was applied on the basis of a series of cognitive labs and unstandardized pre-pilot studies as well as on the basis of standardized pilot ($n = 454$ prospective preschool teachers) and validation studies ($n = 354$ prospective preschool teachers; for results, see Dunekacke, Jenßen & Blömeke, 2015a, b; 2015d, Jenßen, Dunekacke, Eid & Blömeke, 2015) as well as on the basis of conceptual considerations (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014). The three resulting knowledge tests consisted of multiple-choice, bundled, and open-response items. In all cases, gender-neutral language was used to reduce the risk of stereotype threats (Cadinu, Maass, Rosabianca, & Kiesner, 2005) and the language level was kept relatively simple to reduce bias that would favor students at pedagogical colleges.

The assessment of prospective preschool teachers' *MCK* consisted of 24 items that covered the four subdimensions numbers, sets, and operations; shape, space, and change; quantity, measurement, and relations; data, combinatorics, and chance as confirmed by expert validation (Dunekacke, Jenßen & Blömeke, 2015a). Open responses (including drawing figures and finishing tables or formulas) were required for 14 items, whereas 10 were multiple-choice items. These data resulted in 24 dichotomous items that were used to create the *MCK* score.¹ Score reliability was estimated according to Raykov, Dimitrov, and Asparouhov (2010) and was good ($P\gamma = .88$). Figure 2 presents an example item (for scientific purposes, access to the full

¹ Separate scores for each subdimension were not estimated because the distinction between these served conceptual purposes in the context of test development.

instrument can be granted by the first author of this paper; all items were administered in German, those displayed in the following were translated into English for the purposes of this publication).

Insert Figure 2 here

The *MPCK* assessment consisted of 28 items that covered diagnosing children's developmental state in mathematics and designing an informal learning environment that fosters the mathematical learning of children between the ages of 3 and 6 (Dunekacke, Jenßen & Blömeke, 2015b). Open responses were required by five items, whereas 23 were multiple-choice or bundled items. All items were scored dichotomously right or wrong so that the resulting *MPCK* score consisted of 28 items. Score reliability was good ($P\gamma = .87$). Figure 3 presents an example item.

Insert Figure 3 here

The assessment of *GPK* consisted of 18 items that covered general foundations from educational theory, psychology, and instructional research (Blömeke, Jenßen, Dunekacke, Suhl, Grassmann & Wedekind, 2015). Open responses were required by three items, whereas 15 were multiple-choice or bundled items. The information from these items was used to create 18 dichotomous items. Score reliability was lower than for the other two knowledge constructs but still sufficient ($P\gamma = .68$). An example item is displayed in Figure 4.

Psychometric properties of the knowledge tests. To ensure sufficient objectivity in the implementation of the assessments, all procedures such as the timing or use of materials were prescribed in a manual, and administrators of the assessments were trained in a standardized way according to it. Evaluation objectivity was ensured by developing a codebook that described precisely how to code open-ended answers according to their content and which codes to

evaluate as correct (1) or incorrect (0). Interrater reliability was ensured by coding 20% of the open-response items twice, resulting in a good interrater reliability for GPK ($Md_{Kappa} = .76$, Range = .64 to .88; $Md_{Yules} = .98$, Range = .95 to 1.00), MPCK ($Md_{Kappa} = .73$, Range = .64 to .92; $Md_{Yules} = .97$, Range = .92 to 1.00), and MCK ($Md_{Kappa} = .78$, Range = .69 to .86; $Md_{Yules} = .99$, Range = .95 to 1.00; Cohen, 1960; Yules, 1912).

The content validity of the three knowledge tests was confirmed in a standardized procedure by an expert panel. The experts evaluated each single item as well as the entire tests on their representativeness of the respective constructs and their power to predict and explain differences in response behavior of prospective preschool teachers (Jenßen, Dunekacke & Blömeke, 2015).

Factorial validity of inferences drawn from the knowledge test results was confirmed with different samples from the pilot and validation studies by comparing the fit of a three- and a one-dimensional model to the data. The data revealed a better fit of the three-dimensional model. While all three knowledge dimensions were significantly positively related with each other as hypothesized, it was still possible – again as hypothesized – to distinguish them empirically (Jenßen, Dunekacke, Baack, Tengler, Koinzer, Schmude ... Blömeke, 2015).

With the large present sample, it was in addition possible to carry out multiple-group confirmatory factor analysis (MG-CFA; Jöreskog, 1971) which means that each model was estimated in parallel within the four subsamples – prospective preschool teachers at the end and at the beginning of teacher education and this at secondary education and at higher education institutions thus also accounting for the oversampling of the latter. Future teachers represented the first level, and classes represented the second level in these models (two-level CFA mixture

modeling using the known-class and cluster options implemented in MPlus 7.3; Muthén & Muthén, 2014).

Results supported the notion of preschool teachers' knowledge as a three-dimensional construct with latent correlations varying from .62 to .92. The relation between GPK and MCK was the lowest, whereas the strongest relation existed between GPK and MPCK. However, given this strong relation a more parsimonious solution was estimated for exploratory reasons with two dimensions that unified the two latter. This model revealed a similarly good model fit as the three-dimensional model, which indicates that it may not be possible to distinguish GPK and MPCK empirically (see table 3; Blömeke, Jenßen, Dunekacke, Suhl, Grassmann & Wedekind, 2015). Since such a two-dimensional model of preschool teacher knowledge could not yet been replicated with an independent sample and since the OTL provided during teacher education revealed a clear-cut three-dimensional structure (see below), the present study was carried out by applying such three-dimensional models.

Insert Table 3 here

Convergent and discriminant validity of the inferences drawn from the knowledge test results were supported in relation to school marks based on data from the present sample. The better a prospective preschool teacher was in mathematics at school, the higher the teacher's MCK and MPCK scores were ($\beta = .21$ or $\beta = .11$, respectively). By contrast, no significant relation existed between school mathematics and GPK (Blömeke, Jenßen, Dunekacke, Suhl, Grassmann & Wedekind, 2015).

Metric measurement invariance of the three knowledge tests was confirmed across the four different subgroups of the present sample of prospective preschool teacher education students at the beginning or the end of their program at vocational schools or pedagogical

colleges as well as across students of different genders or with different language backgrounds (see Table 4; Blömeke, Jenßen, Dunekacke, Suhl, Grassmann & Wedekind, 2015). This means that it was possible to compare relations between constructs across these groups.

Insert Table 4 here

The criterion validity of inferences drawn from the test results in terms of their relation to performance was supported by data from an earlier sample. MCK and MPCK were direct predictors of prospective preschool teachers' abilities to perceive teacher-children interactions in preschool ($\beta_{\text{direct}} = .45$ or $\beta_{\text{direct}} = .60$, respectively) and were indirect predictors, mediated through perception skills, of their abilities to plan actions that foster children's mathematical development ($\beta_{\text{indirect}} = .43$ or $\beta_{\text{indirect}} = .58$, respectively) demonstrated in a video-based assessment that showed typical preschool situations (Dunekacke, Jenßen & Blömeke, 2015a, b).

As hypothesized on the basis of Ma (1999), an application of latent-state-trait models to an earlier sample revealed that the different subdimensions of MCK and mathematics anxiety were negatively related ($\Psi = -.24$; $\Psi = -.38$; Jenßen, Dunekacke, Eid & Blömeke, 2015). Furthermore, MCK turned out to be stable enough over the course of 3 weeks to be regarded as a trait rather than a state.

Classroom-level predictors. OTL are modeled on the class level because the teacher education institutions are in control of the content and materials they deliver to the prospective preschool teachers. These are assigned to classes that take largely the same OTLs. Neither vocational schools nor pedagogical colleges offer substantial possibilities to choose between content topics.

The prospective preschool teachers reported their OTL in mathematics, mathematics pedagogy, and general pedagogy by rating the coverage of certain topics in each field on 4-point

Likert scales (1 = *not at all*, 4 = *intensely*). The topics were derived from the conceptual framework described in detail in the Appendix. OTL scores represent average item scores so that 1.0 represents the lowest score possible, 4.0 the highest score, and 2.5 the neutral point. Scale reliability for the present sample was evaluated with Cronbach's α , and model fit was evaluated with absolute and relative goodness-of-fit statistics derived from a CFA of the three constructs: OTLs in mathematics, mathematics pedagogy, and general pedagogy (Hu & Bentler, 1999). Comparative Fit Index (CFI) estimates $> .95$ indicate a very good fit, and estimates $> .90$ a good model fit. Root Mean Square Error of Approximation (RMSEA) and Standardized Root Mean Square Residual (SRMR) estimates $< .05$ indicate a very good fit, and estimates $< .08$ a good model fit.

OTL in mathematics were assessed with four items that covered numbers, sets, and operations; shape, space, and change; quantity, measurement, and relations; as well as data, combinatorics, and chance. The scale score's reliability and its model fit were good, $\alpha = .83$; $X^2(2) = 2.85$, $p = .24$; CFI = 1.00; RMSEA = .02, 90% CI [.00, .05], $p = .94$; SRMR = .01. OTLs in mathematics pedagogy were surveyed with seven items that covered the extent to which the prospective preschool teachers had learned to diagnose the mathematical development of children such as their understanding of numbers, shapes, or measurement and to design informal learning environments that foster children's mathematical development in everyday situations or play. The reliability was very good and the model fit was satisfactory, $\alpha = .92$; $X^2(19) = 151.00$, $p < .001$; CFI = 1.00; RMSEA = .06, 90% CI [.05, .07], $p = .02$; SRMR = .04. Finally, the prospective preschool teachers reported their OTL in general pedagogy with four items that covered foundational topics such as basic terms of education and care, teaching methods, or dealing with heterogeneity. The reliability was just satisfactory ($\alpha = .75$), but the model fit was

very good, $\chi^2(5) = 12.79, p = .03$; CFI = 1.00; RMSEA = .03, 90% CI [.01, .05], $p = .96$; SRMR = .05.

On the basis of an MG-CFA (see Table 5), factorial validity with respect to the hypothesized multidimensional structure of OTL was supported by the data. As indicated by the chi-square difference test and the difference in the information criteria reported, the three-dimensional model fit the data significantly better than the one- and the two-dimensional ones. In addition, the factor loadings of all but one item were above the critical threshold of .50 suggested in the literature (Dawis, 2000; Wheaton, 1977), whereas this applied to only half of the items in the one-dimensional and to 11 of the items in the two-dimensional models. Finally, the underlying factor significantly explained variance in the preschool teachers' response behavior of all items in the two- and the three-dimensional models, whereas this did not apply to four items in the one-dimensional model.

Insert Table 5 here

Control variables. All hypotheses were tested by controlling for the individual (level 1) background characteristics of prospective preschool teachers typically found to be predictive of educational outcomes such as gender, family background, and prior knowledge (Blömeke, Suhl, Kaiser & Döhrmann, 2012; Klusmann et al., 2012; Rowan, Correnti, & Miller, 2002; Teddlie & Reynolds, 2004). Gender was coded dichotomously (0 = *female*, 1 = *male*). The language spoken at home as a first indicator of family background was assessed by using a 4-point Likert scale ranging from *never* speaking German at home through *always*. We dichotomized this scale for further analyses into *always* (1) versus the other categories (0) to make up for the skewed distribution (see Table 2). Mother's education as a second indicator of family background was measured on a scale that represented eight educational levels (*below lower secondary* through

PhD). Because of the skewed distribution, it was also dichotomized into a category that included at least a high-school degree (1) versus the lower categories (0). Prior knowledge was assessed with self-reports of participants' most recent marks (class grades based on exams) in school mathematics and German as indicators, which have been shown to be valid indicators of prior knowledge in earlier studies (Blömeke, Suhl, Kaiser & Döhrmann, 2012). All background characteristics were introduced simultaneously on level 1 into the models.

Study Design and Scaling

Participants had 90 min to work on the instruments during the teacher education class they were enrolled in at the time the study was carried out. The instruments were presented in a paper-and-pencil format. We used six test booklets that were randomly distributed in each classroom. Each booklet started with background and OTL variables before the knowledge items followed in a multi-matrix design. The items from each knowledge dimension (i.e., MCK, MPCK, and GPK) were distributed across five blocks (A1, A2, B, C, D) in such a way that each dimension was represented in each block, and the item difficulty of each block was the same on average according to the item difficulty estimates from the pilot studies. The blocks were distributed across the six test booklets so that A1 and A2 were represented in each booklet, whereas B, C, and D were randomly assigned. The blocks were then rotated in such a way that their sequence varied systematically.

To scale the knowledge test data, we applied the so-called Birnbaum model, a 2-parameter logistic item response theory (IRT) model that estimates not only item difficulties but also item discrimination parameters (Andrich, 2004).² The common items from Blocks A1 and A2 served as anchor items. The four subgroups from our sample were defined to have equal

² We refrained from applying a 3-parameter model that would also have included a guessing parameter because sample size was not sufficient to do so and would have resulted in the risk of unstable model estimation.

weights in the scaling process so that the characteristics of one group could not dominate the results. Missing values due to the booklet design of the tests could be regarded as missing completely at random and therefore did not introduce bias (Rubin, 1976). Missing values on items that were skipped or not reached were coded as missing at random (Pohl, Gräfe, & Rose, 2014). The proportion of the two types of missing values was low on all measures with a range of less than 1% to slightly above 4%. These missing values were included in the model estimation in a model-based iterative process by applying the full-information-maximum-likelihood (FIML) method, which uses all information available and is least prone to bias (Lüdtke, Robitzsch, Trautwein, & Köller, 2007). The resulting person estimates were transformed into a mean of 50 and a standard deviation of 10 to facilitate interpretation.

Data Analysis

The data were gathered in a multilevel structure with prospective preschool teachers (individual level, Level 1) nested in teacher education classes (classroom level, Level 2). Therefore, two-level structural equation modeling (SEM) was applied to test all hypotheses except H1d which was tested in a two-level confirmatory factor analysis (CFA) where the fit of models with fixed vs. freely estimated covariances between OTL and knowledge in the four subgroups was compared. The intraclass correlations of .18 (GPK), .19 (MPCK), and .15 (MCK) indicated larger homogeneity within classrooms than if the prospective preschool teachers had been drawn randomly, thus justifying the multi-level approach. Explicitly modeling the cluster structure offers several advantages. First, statistically efficient estimates of regression coefficients and correct standard errors are obtained (Hox, 2002). Second, and this was important in the context of this paper, covariates at any level of the hierarchy can be used, and this makes it possible to examine the extent to which differences in achievement could be predicted by OTL or structural

characteristics of preschool teacher education as class-level variables while controlling for individual preschool teachers' background.

In a few teacher education institutions up to four classes took part in the study which means that there may be some shared variance at the third (school) level. However, sample size was not sufficient to take this third level into account because variance on the second level would not have been sufficient with most institutions participating with only one class. Missing data were handled by applying FIML (Graham, 2003). The individual-level variables were introduced by grand-mean centering (Snijders & Bosker, 2012).

Because the relations between predictors and outcomes might vary across subpopulations, a multiple-group structure was added to the two-level models, which means that all path coefficients were allowed to vary across all groups when testing the hypotheses. *Within* each subpopulation, it could reasonably be assumed that the relations played out in the same way, which means that slopes were fixed across classrooms. We applied a robust maximum-likelihood estimator that could take into account the non-independence of observations due to cluster sampling and would result in the estimation of robust standard errors (Muthén & Muthén, 2014). In all multi-level models, the EAP estimates obtained from the 2PL IRT scaling of MCK, MPCK, and GPK were used as manifest indicators to make these models less complex and identifiable.

In light of the sample size and the moderate complexity of most models, the 1% level of significance was used (with the exception of the more complex mediation model that tested H3 for which the 5% level was used). We report Cohen's d (1988) as the measure of effect sizes with respect to mean differences in predictors and outcomes between the four subgroups examined in this paper. Estimates larger than $d = 0.2$ can be regarded as small, larger than $d =$

0.5 as medium, and larger than $d = 0.8$ as large effects. Differences in regression coefficients were tested for statistical significance based on Clogg, Petkova and Haritou (1995).

Results

Descriptive Statistics

OTL in mathematics and mathematics pedagogy were offered to a lower degree than OTL in general pedagogy, and these findings held at both institutions and at the beginning as well as at the end of teacher education (see Table 6). The differences in OTL reported between the beginning and the end of teacher education had very large effect sizes as indicated by Cohen's d , and this held in all three domains. Prospective preschool teachers at vocational schools reported more OTL than their counterparts at pedagogical colleges at the beginning of their programs, particularly in general pedagogy. By contrast, prospective preschool teachers from pedagogical colleges reported an advantage in OTL in mathematics pedagogy at the end of their programs with a very large effect size.

Insert Table 6 here

GPK and MPCK differed significantly between vocational schools and pedagogical colleges in favor of the latter (see Table 7), and this finding held at the beginning as well as at the end of teacher education. Effect sizes indicating the differences were larger at the end ($d = 0.74$ or $d = 0.88$, respectively) than at the beginning of the programs ($d = 0.31$, $d = 0.50$). Both groups of students had significantly more GPK and MPCK at the end of their teacher education, but again, these differences were larger at pedagogical colleges (around $d = 0.62$) than at vocational schools ($d = 0.21$ or $d = 0.32$, respectively). MCK also differed substantially between vocational schools and pedagogical colleges in favor of the latter (around $d = 0.73$) but not between the beginning and the end of teacher education.

Insert Table 7 here

Relations of Process Characteristics to Prospective Preschool Teachers' Knowledge (H1)

As hypothesized, OTL in *general pedagogy* had a significant relation to GPK (H1a). If prospective preschool teachers reported more OTL in general pedagogy, they scored higher on the GPK assessment, and this occurred at a rate of 4.8 test points for each 1-point increase on the OTL scale (see Table 8). This corresponds to about half of a standard deviation, which is a medium effect size. As hypothesized, the relation between OTL in general pedagogy and GPK was also significantly stronger than the nonsignificant relation between OTL in mathematics and GPK. However, in contrast to our hypothesis the significant relation between OTL in mathematics pedagogy and GPK did not differ significantly from the effect of OTL in general pedagogy. OTL in general pedagogy were not significantly related to MPCK or MCK.

Insert Table 8 here

As hypothesized, OTL in mathematics pedagogy were strongly related to MPCK (H1b). The difference of 5.7 test points for a 1-point increase on the OTL scale represented about half a standard deviation and, thus, a medium effect size. As hypothesized, the relation was significantly stronger than the nonsignificant relation between OTL in mathematics and MPCK whereas it did not differ significantly from the relation between OTL in general pedagogy and MPCK. OTL in mathematics pedagogy were also significantly related to MCK and GPK, each time with small effect sizes of about 4 more test points for a 1-point increase on the OTL scale.

In contrast to H1c, OTL in mathematics did not have a significant relation to MCK. Only OTL in mathematics pedagogy were significantly related to this knowledge dimension.

The relations between OTL and the knowledge indicators played out the same way across all four subgroups, no matter whether the prospective preschool teachers were trained at

vocational schools or in higher education or whether students were at the beginning or the end of their teacher education (H1d; see Table 9). Technically speaking, this means that freeing up the covariances between OTL and knowledge in the respective subgroups did not significantly improve the fit of the multiple-group model; Satorra-Bentler-scaled chi-square difference test (TRd): $\chi^2(8) = 3.7, 9.0, \text{ or } 4.9$, respectively.

Insert Table 9 here

Relations of Structural Preschool Teacher Education Characteristics and Knowledge Outcomes (H2)

The hypothesis that the structural characteristics of preschool teacher education would predict prospective preschool teachers' knowledge was supported by the data (H2). The effect sizes were up to two thirds of a standard deviation (see Table 10). This applied to differential relations of types of institutions (H2a). Prospective preschool teachers from pedagogical colleges achieved significantly higher test scores than students from vocational schools, and this finding held with respect to GPK (+5.4), MPCK (+6.6), and MCK (+7.0).

Insert Table 10 here

The relations of the program stages were significant for GPK and MPCK. The knowledge in these two dimensions was higher at the end than at the beginning of preschool teacher education (H2b). The difference of about 3 test points corresponds to a small effect size. However, in contrast to our hypothesis, MCK was not higher at the end of preschool teacher education than it was at the beginning.

Mediation Models: Relations of Structure and Process to Knowledge (H3)

The relations of the structural preschool teacher education characteristics “type of institution” and “program stage” to GPK, MPCK, and MCK were hypothesized to be at least

partly mediated through the respective domain-specific OTLs (H3). The data supported this hypothesis with respect to program stage and all knowledge indicators as well as with respect to type of institution and MPCK. In contrast, the data did not support this hypothesis with respect to type of institution and GPK or MCK (see Figure 5). These unexpected findings suggest direct paths rather than indirect ones. The individual-level variables of gender, family background, and prior knowledge were controlled for in all three two-level models.

Insert Figure 5 here

With respect to program stage, no significant direct relations to GPK, MPCK, or MCK existed any longer once the respective domain-specific OTL were included. The three knowledge indicators significantly depended on the extent to which OTL were provided in general pedagogy, mathematics pedagogy, or mathematics, which in turn significantly depended on whether the prospective preschool teachers were at the beginning or the end of their training (in favor of the latter). The additional indirect relation of this mediation of structure through process was significant for GPK and MPCK but not for MCK.

The picture differed with respect to the type of institution that provided preschool teacher education. Only in the case of MPCK was the relation of this structural characteristic partly mediated by OTL as a process characteristic. In addition to a significant direct relation of the type of institution to MPCK in favor of pedagogical colleges, a significant indirect relation existed. The OTL score in mathematics pedagogy was significantly related to MPCK, and this score in turn depended significantly on the type of institution (again in favor of pedagogical colleges).

By contrast, the type of institution did not matter significantly for how many OTL were offered in mathematics or in general pedagogy. So, in the cases of MCK and GPK, only

significant direct relations of this structural characteristic to outcomes existed. The fit of all three mediation models was very good (GPK: CFI = .98; RMSEA = .02; SRMR = .01 for the within model and .00 for the between model; MPCK: CFI = 1.00; RMSEA = .00, SRMR = .00 for the within and between models; MCK: CFI = .99; RMSEA = .01; SRMR = .01 for the within model and .00 for the between model).

Discussion

The relation of structural and process teacher education characteristics to outcomes was tested through multilevel modeling with 1,851 prospective preschool teachers nested in 86 classes from vocational school and pedagogical colleges in Germany. Structural and process teacher education characteristics were modeled on the second level (classroom), whereas teacher background was controlled for on the first level (individual). The three knowledge indicators were modeled on both levels with data gathered through standardized and domain-specific testing, thus closing a much criticized gap in preschool research (Early et al., 2007; Whitebook et al., 2009).

The descriptive statistics revealed that OTL in mathematics and mathematics pedagogy were offered less often during preschool teacher education than OTL in general pedagogy confirming that the traditional concept of stronger emphasis on care than on cognitive development (Liegler, 2008) is still shaping preschool teacher education in Germany (for similar results in the US see Isenberg, 2000). This applied interestingly to programs at both types of institutions: vocational schools on the (post)secondary level and pedagogical colleges on the tertiary level – a result that may demonstrate that moving preschool teacher education up to the tertiary level alone may not be sufficient to change its nature. Prospective teachers from both types of institutions reported more OTL in general pedagogy already when they entered teacher

education. This may go back to entrance requirements: the completion of vocational training in a care profession for vocational school students or a pedagogical internship for college students.

The data supported most but not all of our hypotheses. OTL in general pedagogy (H1a) and mathematics pedagogy (H1b) as well as types of institutions (H2a) and program stage (H2b) had significant relations to GPK or MPCK, respectively. Berliner's (1995) early call for a domain-specific perspective on teaching and learning contexts was thereby for the first time supported with respect to preschool teacher education in Germany.

The data also supported H3 that the effects of the distal structural preschool teacher education characteristics "type of institution" and "program stage" were partly mediated by OTL as proximal process characteristics. This applied in particular to program stage, and in the case of MPCK, also to the type of institution. Process characteristics are obviously as crucial in the development of knowledge during preschool teacher education as structural characteristics. This result opens up for interesting conclusions beyond just moving programs from the secondary to the tertiary level (see below).

Outcomes of preschool teacher education in terms of GPK and MPCK already differed significantly at the beginning of the programs between vocational schools and pedagogical colleges in favor of the latter, which is probably an indicator of the stronger school credentials required by pedagogical colleges (graduation from high school instead of middle school). The differences between the beginning and the end of preschool teacher education were larger at pedagogical colleges than at vocational schools, a finding that can be interpreted as an indication of more OTL delivered during the longer programs at pedagogical colleges. Both results, the differences at the beginning and the differential development during teacher education shed light on the much under-researched preschool teacher education below the tertiary level (Waller,

2006). They point to severe disadvantages of this group of prospective teachers compared to college students.

OTL in mathematics and MCK behaved differently compared to MPCK and GPK. In contrast to our hypotheses H1c and H2b, neither a significant difference in MCK existed between the beginning and the end of teacher education, indicating a lack of progress during the programs, nor was there a significant relation between OTL in mathematics and MCK. Only OTL in mathematics pedagogy were significantly related to MCK. It seems as if OTL in mathematics pedagogy were better able to support MCK development although the correlational nature of the data asks for caution here.

Given that the content validity of the MCK test was confirmed in standardized expert reviews (Jenßen, Dunekacke & Blömeke, 2015), the learning of MCK during preschool teacher education needs more research. It may be the case that the high degree of math anxiety found in a different sample of prospective preschool teachers played out negatively (Jenßen, Dunekacke, Eid & Blömeke, 2015) or that the more applied nature of mathematics in mathematics pedagogy OTL facilitated the acquisition of MCK for this group of teachers. If the latter result can be replicated, it has implications for preschool teacher education design and further research. In contrast to primary or secondary teachers where OTL in mathematics played a crucial role and MCK turned out to be a necessary prerequisite for MPCK (Blömeke, Suhl, Kaiser & Döhrmann, 2012; Tatto et al., 2012), the findings of the present study suggest that OTL in mathematics pedagogy may be more beneficial for acquiring MCK rather than OTL in mathematics.

This would be a unique result that distinguishes prospective preschool teachers from others. MPCK may build more appropriately on preschool teachers' prior knowledge because students come into the programs with more experience in general pedagogy. Learning typically

happens through connecting new information to prior knowledge (Carroll, 1963). OTL in mathematics pedagogy may elicit teachers' prior knowledge and help them to make connections to new knowledge, thus serving as a bridge and therefore being an effective instructional approach.

Before conclusions are drawn from these results and interpretations, some methodological limitations of the study need to be pointed out. Institutions provide a set of intertwined organizational and pedagogical characteristics (Tinto, 1998) so that other characteristics than those examined here could be causal (e.g., climate or composition effects). It is not possible to disentangle such effects in a correlational study. Furthermore, the fuzzy notion of “OTL” which has been defined in different ways in educational research, leads to a lot of noise in the state of research. The results would therefore be substantially strengthened if they were replicated with other samples. Other researchers in the field are requested to apply the OTL measures and knowledge tests in further samples, in particular in other countries, or to develop new measures based on the same construct definitions so that cross-validations can take place. This is necessary to avoid acting too quickly on the basis of one study that was conducted in only one national context.

Conclusions

Given the small amount of OTL offered in mathematics pedagogy as indicated by the descriptive statistics, the findings presented in this paper have to be of concern with respect to fostering children's mathematics literacy (Duncan et al., 2006; Reynolds, 1995). Even though preschool standards require preschool teachers to achieve ambitious objectives (Jenßen, Dunekacke & Blömeke, 2015), prospective teachers do not seem to receive sufficient OTL. Because of the interdisciplinary nature of MPCK as an amalgam of MCK and GPK (Shulman,

1986), this dimension of knowledge seems to be crucial for the development of teacher knowledge in general. The significant relation between OTL in mathematics pedagogy and outcomes indicates that it may be worthwhile to increase the number of OTL in this domain, particularly at vocational schools (Janssen, 2010).

Note that the conclusion that more OTLs should be provided differs from a request to train all preschool teachers at the higher education level. The finding that processes mediate structural characteristics may open up new ways of thinking in this context. If it is OTL that count, these can also be provided at vocational schools. However, differential intake characteristics have still to be taken into account. Furthermore, the colleges have the advantage that their preschool teacher education program lasts for 4 years instead of only 2, thus providing more teaching time to prepare preschool teachers for all the tasks they have to cover. The question of how to provide the higher costs coming with longer preschool teacher education has to be taken into account in this context, too.

Besides such policy-related conclusions, a broad range of research needs can be derived from the present study. Future studies should address mediations through cross-domain OTL in particular with respect to mathematics pedagogy and general pedagogy. Furthermore, it needs to be examined how preschool teachers knowledge base is transformed into job performance because ultimately preschool teacher education is designed to support high quality teaching in preschool and child development. Not many studies have taken on examining such long-term effects of teacher education because this needs sophisticated designs and standardized observation protocols or video-based measures of teaching skills. Although difficult to implement such research would provide urgently needed information on how to structure

preschool teacher education so that prospective teachers acquire an appropriate knowledge and skill base to succeed in their job.

References

- Abell Foundation (2001). *Teacher certification reconsidered: Stumbling for quality*. Baltimore: Abell.
- American Educational Research Association, American Psychological Association & National Council on Measurement in Education (2014). *The Standards for Educational and Psychological Testing*. Washington, D.C.: AERA.
- Anders, Y. (2012). *Modelle professioneller Kompetenzen für fröhpädagogische Fachkräfte: Aktueller Stand und ihr Bezug zur Professionalisierung*. München, Germany: Knoblingesign.
- Andrich, D. (2004). Controversy and the Rasch model: A characteristic of incompatible paradigms? *Medical Care*, 42, 1–16.
- Blömeke, S., Kaiser, G. & Lehmann, R. (Eds.) (2008). *Professionelle Kompetenz angehender Lehrerinnen und Lehrer: Wissen, Überzeugungen und Lerngelegenheiten deutscher Mathematikstudierender und -referendare – Erste Ergebnisse zur Wirksamkeit der Lehrerausbildung*. Münster, Germany: Waxmann.
- Blömeke, S., Suhl, U. & Kaiser, G. (2011). Teacher education effectiveness: Quality and equity of future primary teachers' mathematics and mathematics pedagogical content knowledge. *Journal of Teacher Education*, 62, 154-171.
- Blömeke, S., Suhl, U., Kaiser, G. & Döhrmann, M. (2012). Family background, entry selectivity and opportunities to learn: What matters in primary teacher education? An international comparison of fifteen countries. *Teaching and Teacher Education*, 28, 44-55.
- Blömeke, S., Jenßen, L., Dunekacke, S., Suhl, U., Grassmann, M. & Wedekind, H. (2015). Professionelle Kompetenz von Erzieherinnen messen: Entwicklung und Validierung standardisierter Leistungstests für fröhpädagogische Fachkräfte. *Zeitschrift für Pädagogische Psychologie*, 29, 177-191.
- Bassok, D. (2012). *Raising teacher education levels in Head Start: Are there program-level tradeoffs?* (= CEPWC Working Paper; 3). Charlottesville, VA: Center on Education Policy and Workforce Competitiveness.
- Berliner, D. C. (1985). Effective classroom teaching: the necessary but not sufficient condition for developing exemplary schools (pp. 127-154). In G. R. Austin & H. Garber (Eds.), *Research on exemplary schools*. Orlando, Fl.: Academic Press Ins.
- Bogard, K., Traylor, F. & Takanishi, R. (2008). Teacher education and PK outcomes: Are we asking the right questions? *Early Childhood Research Quarterly*, 23, 1–6.

- Bryant, F. B. & Satorra, A. (2012). Principles and practice of scaled difference chi-square testing. *Structural Equation Modeling*, *19*, 372-398.
- Burchinal, M. R., Cryer, D., Clifford, R. M. & Howes, C. (2002). Caregiver training and classroom quality in child care centers. *Applied Developmental Science*, *6*, 2-11.
- Carroll, J. B. (1963) A model of school learning. *Teachers College Record*, *64*, 722-733.
- Clements, D. H., Sarama, J., & DiBiase, A.-M. (Eds.). (2004). *Engaging young children in mathematics: Standards for early childhood mathematics*. Mahwah, NJ: Lawrence Erlbaum.
- Clogg, C. C., Petkova, E., & Haritou, A. (1995). Statistical methods for comparing regression coefficients between models. *American Journal of Sociology*, *100*, 1261-1293.
- Cohen, J. (1960). A coefficient for agreement for nominal scales. *Education and Psychological Measurement*, *20*, 37-46.
- Connor, C. M., Morrison, F. J. & Slominski, L. (2006). Preschool instruction and children's emergent literacy growth. *Journal of Educational Psychology*, *98*, 665-689.
- Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives*, *8*(1). Retrieved from <http://epaa.asu.edu/ojs/article/view/392/515>.
- Dawis, R. V. (2000). Scale construction and psychometric considerations. In H. E. A. Tinsley & S. D. Brown (Eds.), *Handbook of applied multivariate statistics and mathematical modeling* (pp. 65-94). Orlando, FL: Academic Press.
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P. ... Japel, C. (2006). School Readiness and Later Achievement. *Developmental Psychology*, *43*, 1428-1462.
- Dunekacke, S., Jenßen, L. & Blömeke, S. (2015a). Effects of mathematics content knowledge on pre-school teachers' performance: A video-based assessment of perception and planning abilities in informal learning situations. *International Journal of Science and Mathematics Education*, *13*, 267-286.
- Dunekacke, S., Jenßen, L. & Blömeke, S. (2015b). Mathematikdidaktische Kompetenz von Erzieherinnen und Erziehern: Validierung des KomMa-Leistungstests durch die videogestützte Erhebung von Performanz. *Zeitschrift für Pädagogik*; *61. Beiheft*, 80-99.

- Early, D. M. & Winton, P. J. (2001). Preparing the workforce: Early childhood teacher preparation at 2- and 4-year institutions of higher education. *Early Childhood Research Quarterly, 16*(3), 285-306.
- Early, D. M., Bryant, D. M., Pianta, R. C., Clifford, R. M., Burchinal, M. R., Ritchie, S., et al. (2006). Are teachers' education, major, and credentials related to classroom quality and children's academic gains in pre-kindergarten? *Early Childhood Research Quarterly, 21*, 174–195.
- Early, D. M., Maxwell, K. L., Burchinal, M., Bender, R. H., Ebanks, C., Henry, G. T., et al. (2007). Teachers' education, classroom quality, and young children's academic skills: Results from seven studies of preschool programs. *Child Development, 78*, 558–580.
- Graham, J. W. (2003). Adding missing-data-relevant variables to FIML-based structural equation models. *Structural Equation Modeling, 10*, 80-100.
- Hamre, B. K., Pianta, R. C., Burchinal, M., Field, S., LoCasale-Crouch, J., Downer, J. T., ... Scott-Little, C. (2012). A course on effective teacher-child interactions: Effects on teacher beliefs, knowledge, and observed practice. *American Educational Research Journal, 49*(1), 88-123.
- Howes, C., Whitebook, M., & Phillips, D. (1992). Teacher characteristics and effective teaching in child care: Findings from the National Child Care Staffing Study. *Child & Youth Care Forum, 21*, 399-414.
- Hox, J. (2002). *Multilevel analysis: Techniques and applications*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Hu, L. T. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55.
- Isenberg, J. P. (2000). The state of the art in early childhood professional preparation. In *New teachers for a new century: The future of early childhood professional preparation* (pp. 17-58). Washington, DC: National Institute on Early Childhood Development and Education, U.S. Department of Education.
- Janssen, R. (2010). *Die Ausbildung Frühpädagogischer Fachkräfte an Berufsfachschulen und Fachschulen: Eine Analyse im Ländervergleich*. Munich, Germany: DJI.

- Jenßen, L., Dunekacke, S. & Blömeke, S. (2015). Qualitätssicherung in der Kompetenzforschung: Empfehlungen für den Nachweis von Validität in Testentwicklung und Veröffentlichungspraxis. *Zeitschrift für Pädagogik*; 61. Beiheft, 11-31.
- Jenßen, L., Dunekacke, S., Eid, M. & Blömeke, S. (2015). The relationship of mathematical competence and mathematics anxiety: An application of latent state-trait theory. *Zeitschrift für Psychologie*, 223, 31–38.
- Jenßen, L., Dunekacke, S., Baack, W., Tengler, M., Koinzer, T., Schmude, C., Wedekind, H., Grassmann, M. & Blömeke, S. (2015). KomMa: Kompetenzmodellierung und Kompetenzmessung bei frühpädagogischen Fachkräften im Bereich Mathematik (pp. 59-79). In B. Koch-Priewe, A. Köker, J. Seifried & E. Wuttke (Eds.), *Kompetenzerwerb an Hochschulen: Modellierung und Messung. Zur Professionalisierung angehender Lehrerinnen und Lehrer sowie frühpädagogischer Fachkräfte*. Bad Heilbrunn, Germany: Klinkhardt.
- Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika*, 36, 409–426.
- Kleickmann, T., Richter, D., Kunter, M., Elsner, J., Besser, M., Krauss, S., & Baumert, J. (2013). Pedagogical content knowledge and content knowledge: The role of structural differences in teacher education. *Journal of Teacher Education*, 64, 90-106.
- Klusmann, U., Kunter, M., Voss, T. & Baumert, J. (2012). Berufliche Beanspruchung angehender Lehrkräfte: Die Effekte von Persönlichkeit, pädagogischer Vorerfahrung und professioneller Kompetenz. *Zeitschrift für Pädagogische Psychologie*, 26, 275-290.
- König, J., Blömeke, S., Paine, L., Schmidt, W. H. & 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, 188-201.
- Landry, S. H., Anthony, J. L., Swank, P. R. & Monseque-Bailey, P. (2009). Effectiveness of comprehensive professional development for teachers of at-risk preschoolers. *Journal of Educational Psychology*, 101, 448–465.
- Liegle, L. (2008). Erziehung als Aufforderung zur Bildung: Aufgaben der Fachkräfte in Tageseinrichtungen für Kinder in der Perspektive der frühpädagogischen Didaktik. In W. Thole, H.-G. Roßbach, M. Fölling-Albers & R. Tippelt (Eds.), *Bildung und Kindheit. Pädagogik der Frühen Kindheit in Wissenschaft und Lehre* (pp. 85–113). Opladen: Budrich.

- Lobman, C., Ryan, Sh. & McLaughlin, J. (2005). Reconstructing teacher education to prepare qualified preschool teachers: Lessons from New Jersey. *Early Childhood Research & Practice*, 7(2).
- Ma, X. (1999). A meta-analysis of the relationship between anxiety toward mathematics and achievement in mathematics. *Journal for Research in Mathematics Education*, 30, 520–540.
- Muthén, L. K. & Muthén, B. O. (2014). *Mplus User's Guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.
- National Research Council (2009). *Mathematics learning in early childhood: Paths toward excellence and equity*. Washington, DC: National Academies Press.
- OECD (2004). *Starting strong: Curricula and pedagogies in early childhood education and care*. Paris: OECD.
- Pianta, R. C., DeCoster, J., Cabell, S., Burchinal, M., Hamre, B. K., Downer, J., ... A. Howes, C. (2014). Dose–response relations between preschool teachers' exposure to components of professional development and increases in quality of their interactions with children. *Early Childhood Research Quarterly*, 29, 499–508.
- Piasta, S., Logan, J. A. R., Yeager Pelatti, Ch., Capps, J. L. & Petrill, St. A. (2015). Professional development for early childhood educators: Efforts to improve math and science learning opportunities in early childhood classrooms. *Journal of Educational Psychology*, 107(2), 407–422.
- Pohl, S., Gräfe, L. & Rose, N. (2014). Dealing with omitted and not-reached items in competence tests. Evaluating Approaches Accounting for Missing Responses in Item Response Theory Models. *Educational and Psychological Measurement*, XX(X) 1–30.
- Reynolds, A. (1995). One year of preschool intervention or two: Does it matter? *Early Childhood Research Quarterly*, 10, 1–31.
- Rowan, B., Correnti, R., & Miller, R. J. (2002). *What large-scale, survey research tells us about teacher effects on student achievement: Insights from the Prospects Study of Elementary Schools* (= CPRE Research Report Series RR-051). Pennsylvania: University of Pennsylvania.
- Rubin, D. (1976). Inference and missing data. *Biometrika*, 63, 581–592.
- Scheerens, J. & Blömeke, S. (2016). Integrating teacher education effectiveness research into educational effectiveness models. *Educational Research Review*, 18, 70–87.

- Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15, 4–14.
- Snijders, T. A. B. & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). London: Sage.
- Statistisches Bundesamt (2010). *Frauen und Männer in verschiedenen Lebensphasen*. Wiesbaden: Statistisches Bundesamt.
- Statistisches Bundesamt (2014). *Statistiken der Kinder- und Jugendhilfe: Kinder und tätige Personen in Tageseinrichtungen und in öffentlich geförderter Kindertagespflege am 01.03.2014*. Wiesbaden: Statistisches Bundesamt.
- Tatto, M. T., Schwille, J., Senk, S. L., Ingvarson, L., Rowley, G., Peck, R., et al. (2012). *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries: Findings from the IEA Teacher Education and development Study in Mathematics (TEDS-M)*. Amsterdam: IEA.
- Teddlie, C. & Reynolds, D. (2000). *The international handbook of school effectiveness research*. London: Falmer Press.
- Tinto, V. (1998). Colleges as communities: Taking research on student persistence seriously. *The Review of Higher Education*, 21, 167-177.
- Tout, K., Zaslow, M. & Berry, D. (2005). Quality and qualifications: Links between professional development and quality in early care and education settings. In M. Zaslow & I. Martinez-Beck (Eds.), *Critical issues in early childhood professional development* (pp. 77-110). Baltimore: Paul H. Brooks.
- Travers, K. J. & Westbury, I. (1989). *The IEA study of mathematics I: Analysis of mathematics curricula*. Oxford: Pergamon Press.
- Voss, Th., Kunter, M. & Baumert, J. (2011). Assessing teacher candidates' general pedagogical/psychological knowledge: Test construction and validation. *Journal of Educational Psychology*, 103(4), 952-969.
- Wallet, P. (2006). *Pre-Primary teachers: A global analysis of several key education indicators*. Paper commissioned for the EFA Global Monitoring Report 2007 “Strong foundations: early childhood care and education” (2007/ED/EFA/MRT/PI/32). Paris: UNESCO.
- Weinert, F. E. (2001). Concept of Competence: A Conceptual Classification. In D. S. Rychen & L. Hersh Salganik (Eds.), *Defining and Selecting Key Competencies*. Göttingen: Hogrefe.

Wheaton, B., Muthén, B., Alwin, D. F. & Summers, G. F. (1977). Assessing reliability and stability in panel models (pp. 84–136). In D. R. Heise (Ed.), *Sociological Methodology*. San Francisco: Jossey-Bass.

Whitebook, M., Gomby, D., Bellm, D., Sakai, L. & Kipris, F. (2009). *Preparing teachers of young children: The current state of knowledge, and a blueprint for the future*. Berkeley, CA: Center for the Study of Child Care Employment.

Wilson, M. (2005). *Constructing Measures: An Item Response Modeling Approach*. Mahwah, NJ: Lawrence Erlbaum Associates.

Yule, G. U. (1912). On the methods of measuring association between two attributes. *Journal of the Royal Statistical Society*, 75, 579–642.

Table 1

Sample Size

Type of institution	Vocational school	Pedagogical college	Overall
Program stage	(secondary education)	(higher education)	
First year	594 (32%)	287 (15%)	881 (47%)
Last year	774 (42%)	196 (11%)	970 (53%)
Overall	1.368 (74%)	483 (26%)	1.851 (100%)

Table 2

Descriptive Statistics of the Sample by Subgroup

	First year vocational school	Last year vocational school	First year pedagogical college	Last year pedagogical college
Age in years (Range)	22 (17-53)	24 (18-54)	22 (18-47)	26 (19-53)
Gender (female)	85%	83%	90%	90%
German language background (always spoken at home)	88%	89%	83%	86%
No. of books at home (> 200)	23%	24%	41%	44%
Mother's education (at least a high-school degree)	17%	16%	32%	27%
Participant's own education (at least a high-school degree)	36%	44%	99%	99%
No. of years of mathematics in school (≤ 10)	47%	48%	2%	5%

Table 3.

Fit of one-, two- and three-dimensional models of prospective teachers' knowledge

Modell	# Par	<i>LL</i>	<i>SCF</i>	<i>AIC</i>	<i>BIC_{adj}</i>
One-dimensional	143	-42535.4	1.92	85356.9	85692.4
Three-dimensional	146	-42398.1	1.90	85088.3	85430.8
Two-dimensional	144	-42403.3	1.90	85094.7	85432.6

Note. # Par=no. parameters, LL=loglikelihood, SCF=Scaling Correction Factor, AIC=Akaike's Information Criterion, *BIC_{adj}*=adjusted Bayesian Information Criterion.

Table 4.

Testing of measurement invariance across different subgroups

Model	no. parameters			log likelihood			model comparison									
	1	2	3	1	2	3	chi-square			df			p value			
Configural	575	287	287	-41 894.0	-40 485.8	-40 528.0	---	---	---	---	---	---	---	---	---	---
Metric	374	220	220	-41 999.2	-40 514.9	-40 567.8	200.7	50.5	48.2	201	67	67	ns	ns	ns	ns
Scalar	173	153	153	-42 216.5	-40 584.4	-40 617.1	381.2	138.1	97.9	201	67	67	<.05	<.05	<.05	<.05

Note. df=degrees of freedom, ns=non significant, subgroups: Model 1=4 groups (beginning or end of teacher education at vocational schools or colleges of education), model 2=2 groups (male, female teachers), model 3=two groups (language always German, not always German), model comparisons: metric vs. configural, scalar vs. metric.

Table 5

Multiple-Group Confirmatory Factor Analysis

OTL model	Loglikeli- hood	# of par.	AIC	BIC	BIC _{adj}	# items a)	# items b)
One- dimensional	-33,595.5	66	67,322.9	67,687.1	67,477.4	7	4
Two- dimensional	-32,471.6	71	65,085.3	65,477.1	65,251.5	11	0
Three- dimensional	-31,854.7	78	63,865.5	64,295.9	64,048.1	14	0

Note. OTL = opportunities to learn, AIC = Akaike information criterion, BIC = Bayesian information criterion, BIC_{adj} = sample-size adjusted Bayesian information criterion, # items a) = # of items with standardized factor loadings $\geq .50$, # items b) = # of items where the variance was not significantly explained by the underlying latent variable. Chi-square difference test: $X^2(5)_{1_2} = 1,123.9, p < .001$; $X^2(7)_{2_3} = 616.9, p < .001$; $X^2(12)_{1_3} = 1,740.8, p < .001$.

Table 6

OTL in General Pedagogy (1), Mathematics Pedagogy (2), and Mathematics (3) Provided during Preschool Teacher Education

	Vocational school			Pedagogical college			<i>t</i> est. for diff. between types of institutions			Cohen's <i>d</i>		
	1	2	3	1	2	3	1	2	3	1	2	3
First year	2.91	1.86	1.81	2.67	1.79	1.66	-9.0*	-2.0 ^{ns}	-4.5*	0.64	0.14	0.30
<i>M</i> (<i>SD</i>)	(.34)	(.27)	(.40)	(.41)	(.68)	(.61)						
Last year	3.27	2.30	2.29	3.22	2.76	2.27	-3.3*	13.8*	-.82 ^{ns}	0.26	1.12	0.08
<i>M</i> (<i>SD</i>)	(.18)	(.42)	(.42)	(.20)	(.40)	(.32)						
<i>t</i> est. for diff. between program stages	25.4*	22.3*	21.5*	17.2*	17.9*	12.8*						
Cohen's <i>d</i>	1.33	1.28	1.17	1.80	1.80	1.31						

Note. OTL = opportunities to learn; *M* = Mean, *SD* = Standard Deviation. Estimates were based on *t* tests of mean differences for independent samples.

* $p < .01$.

Table 7

Prospective Preschool Teachers' GPK (1), MPCK (2), and MCK (3) Scores

		Vocational school			Pedagogical college			<i>t</i> est. for diff. between types of institutions			Cohen's <i>d</i>		
		1	2	3	1	2	3	1	2	3	1	2	3
First	<i>M (SD)</i>	48 (9.8)	47	48	51	52	55	4.6*	6.8*	10.4*	0.31	0.50	0.74
year			(9.9)	(9.5)	(9.9)	(10.0)	(9.3)						
Last	<i>M (SD)</i>	50 (9.5)	50	48	57	58	55	9.7*	11.1*	8.8*	0.74	0.88	0.73
year			(9.2)	(9.6)	(9.2)	(8.7)	(9.5)						
Com-	<i>t</i> est. for diff.	4.4*	5.7*	0.8	7.2*	6.9*	0.2						
parison	between program												
	stages												
	Cohen's <i>d</i>	0.21	0.32	0.00	0.62	0.63	0.00						

Note. *M* = Mean, *SD* = Standard Deviation of the person estimates from the 2PL IRT scaling; Estimates were based on *t* tests of mean differences for independent samples.

* *p* < .01.

Table 8

Two-Level Models of Relations of OTL Provided during Preschool Teacher Education to Prospective Preschool Teachers' Knowledge (in test points)

	GPK	MPCK	MCK
<i>Predictors</i>			
OTL in general pedagogy	+4.8*	+1.8	-2.5
OTL in mathematics pedagogy	+3.7*	+5.7*	+4.3*
OTL mathematics	-2.5	-1.8	-1.3

Note. GPK = General pedagogical knowledge, MPCK = Mathematics pedagogical content knowledge, MCK = Mathematics content knowledge, OTLs = Opportunities to learn. In all models, preschool teachers' gender, school marks in mathematics and German, language background, and mother's education were controlled for on the individual level.

* $p < .01$.

Table 9

Fit of Models for Testing whether Relations of OTL to Outcomes were Different in the Subgroups

Model	GPK			MPCK			MCK		
	LL	# par	BIC _{adj}	LL	# par	BIC _{adj}	LL	# par	BIC _{adj}
Equal	-7,257.8	13	14,571.2	-7,244.0	13	14,543.6	-7,221.2	13	14,498.0
Free	-7,254.1	21	14,597.9	-7,235.0	21	14,559.8	-7,216.3	21	14,522.5

Note. BIC_{adj} = adjusted Bayesian information criterion, LL = Loglikelihood. Equal = covariances constrained to be equal in the four subgroups, free = covariances estimated freely.

Table 10

Two-Level Models of Structural Preschool Teacher Education Effects on Prospective Preschool Teachers' Knowledge (in test points)

	GPK	MPCK	MCK
<i>Predictors</i>			
Type of institution	+5.4*	+6.6*	+7.0*
Program stage	+3.0*	+2.9*	-0.2

Note. GPK = General pedagogical knowledge, MPCK = Mathematics pedagogical content knowledge, MCK = Mathematics content knowledge; type of institution: 0 = vocational school, 1 = pedagogical college, program stage: 0 = beginning of teacher education, 1 = end of teacher education. In both models, preschool teachers' gender, school marks in mathematics and German, language background, and mother's education were controlled for on the individual level.

* $p < .01$.

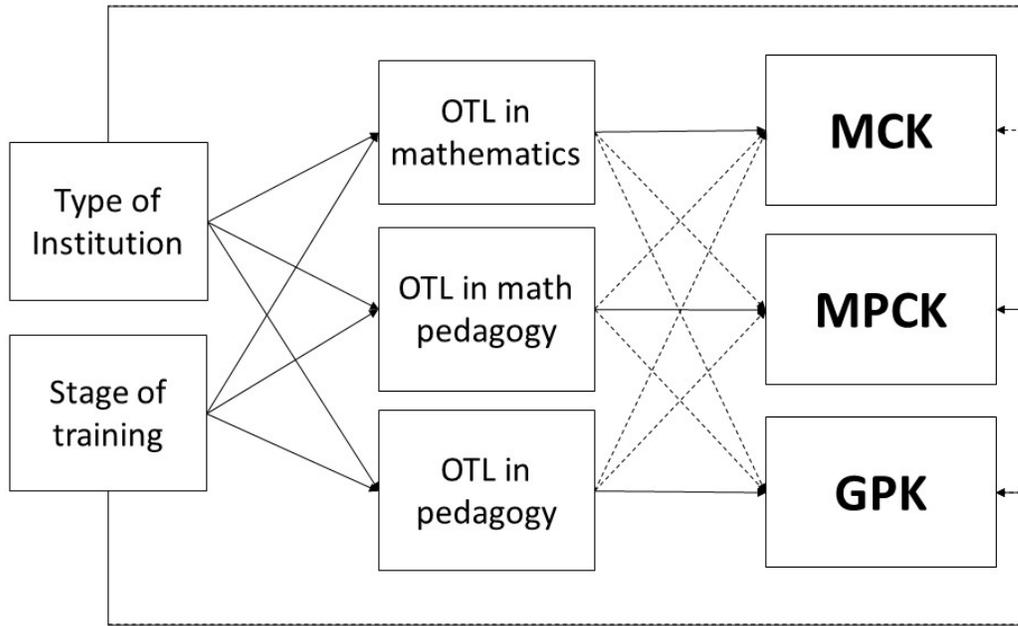


Figure 1. Hypothesized research model (OTL = opportunities to learn, MCK = mathematics content knowledge, MPCK = mathematics pedagogical content knowledge, GPK = general pedagogical knowledge; dotted lines were hypothesized to be weaker than solid lines).



Chris has a blue, a green, a red and a yellow cube.
Chris wants to pile up a tower with the four cubes

Which arithmetic expression provides the number of possibilities of the different towers?

Please indicate your answer with a cross

- $4 + 4 + 4 + 4$
- $4 \cdot 4 \cdot 4 \cdot 4$
- $4 \cdot 3 \cdot 2$
- $4 + 3 + 2$

Figure 2. Example item from the MCK test (translated).

You are playing a dice game with three children. *Please explain, in short, why their mathematical learning in the following field is fostered:*

Numbers and operations (e.g., calculating):

Figure 3. Example item from the MPCK test (subdomains data and modelling; translated).

Some children in your group are playing a strategy game. When they are done, you talk to those who lost and you inquire about their reasoning about why they lost.

Child A: "I was just unlucky."

Child B: "I was not that interested in the game."

Child C: "I do not understand this type of game."

Which child provides a reason that is particularly unfavorable from a motivational point of view? Child _____

Figure 4. Example item from the GPK assessment (translated).

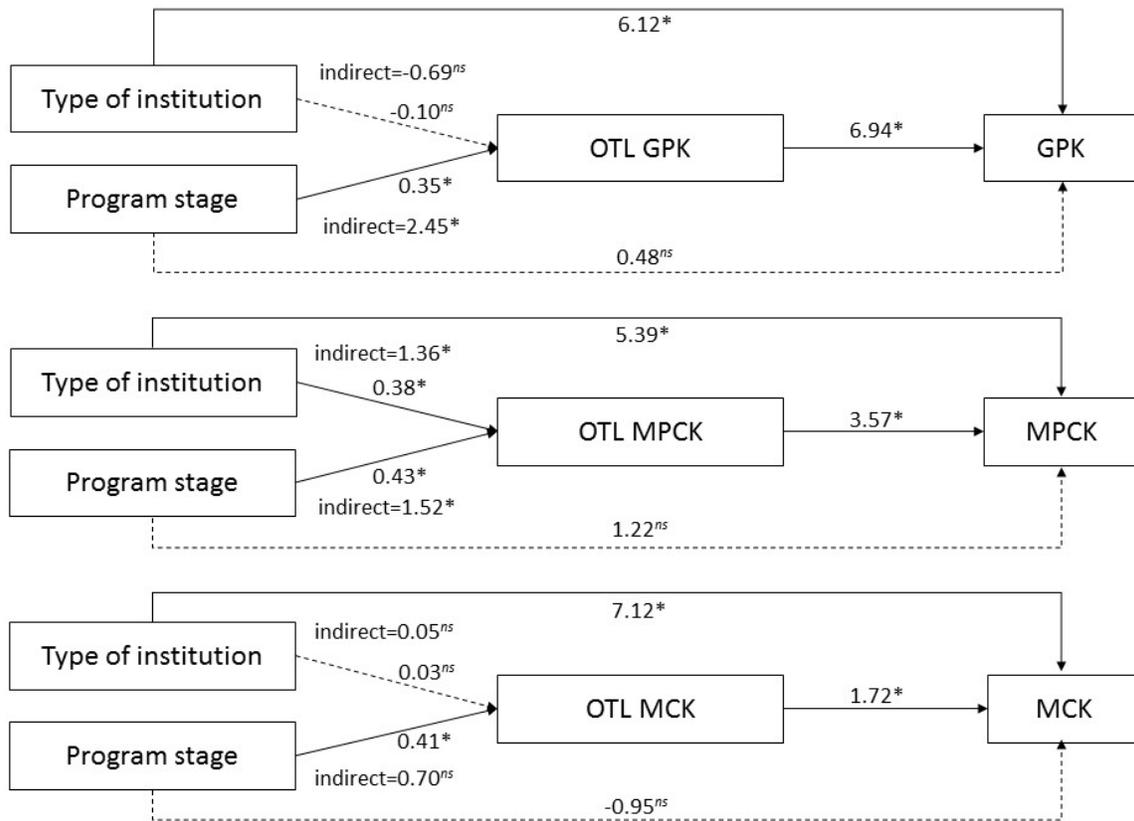


Figure 5. Mediation models of the relations between structural preschool teacher education characteristics, process characteristics, and outcomes (individual-level background variables are controlled for).

Appendix

Dimensions, subdimensions, and descriptors of preschool teacher knowledge (in parentheses: no. of test items)

General pedagogical knowledge (18)	Mathematics pedagogical content knowledge (28)	Mathematics content knowledge (24)
<p>Educational foundations (5)</p> <ul style="list-style-type: none"> - Knowledge of fundamental educational terms (1) - Selection of educational objectives for children aged 3-6 (1) - Application of educational approaches (2) - Formal and informal opportunities to learn (1) <p>Psychological foundations (6)</p> <ul style="list-style-type: none"> - Knowledge of motivation and attribution theories (2) - Diagnosing general learning and developmental processes of 3-to-6-year-old children (2) - Development of strategies to change child behavior (2) <p>Instructional foundations (7)</p> <ul style="list-style-type: none"> - Application of communication and collaboration approaches (2) - Application of approaches to foster learning and development in heterogeneous groups of children between the ages of 3 and 6 (3) - Application of inclusive principles (2) 	<p>Diagnosing children's mathematical development (17)</p> <ul style="list-style-type: none"> - Developmental psychology of children's mathematical competencies (2) - Diagnosing developmental states in the field of number, sets, and operations based on children's statements (5) - Diagnosing developmental states in the field of shape, space, and change based on children's statements (2) - Evaluation of standardized and unstandardized diagnostic approaches (2) - Identification of everyday-life situations with relations to numbers, sets, and operations (3) - Identification of everyday-life situations with relations to shape, space, and change (1) - Identification of everyday-life situations with relations to quantity, measurement, and relations (2) <p>Designing informal learning environments that foster mathematical learning (11)</p> <ul style="list-style-type: none"> - Application of approaches that support mathematical learning (incl. specifics for children at risk) (3) - Initiate play-based experiences with numbers, sets, and operations (2) - Initiate play-based experiences with shape, space, and change (4) - Initiate play-based experiences with data, combinatorics, and chance (1) - Initiate play-based experiences with quantity, measurement, and relations (1) 	<p>Numbers, sets, and operations (6)</p> <ul style="list-style-type: none"> - Knowledge of number range (1) - Application of basic operations (2) - Application of number principles (2) - Understanding sets (1) <p>Shape, space, and change (6)</p> <ul style="list-style-type: none"> - Application of formulas (2) - Recognizing geometrical shapes (2) - Demonstrating space orientation (1) - Constructing geometrical shapes (1) <p>Data, combinatorics, and chance (6)</p> <ul style="list-style-type: none"> - Generating tables and lists of frequencies (2) - Estimating the number of possibilities (1) - Estimating chance (3) <p>Quantity, measurement, and relations (6)</p> <ul style="list-style-type: none"> - Relating speed to time (2) - Transforming verbal into mathematical statements and vice versa (2) - Pattern recognition (2)