



HAL
open science

Relation of a test for assessing pre-service primary teachers' pedagogical content knowledge to other educational variables

Malte Lehmann, Bettina Roesken-Winter, Lars Jenßen

► To cite this version:

Malte Lehmann, Bettina Roesken-Winter, Lars Jenßen. Relation of a test for assessing pre-service primary teachers' pedagogical content knowledge to other educational variables. Thirteenth Congress of the European Society for Research in Mathematics Education (CERME13), Alfréd Rényi Institute of Mathematics; Eötvös Loránd University of Budapest, Jul 2023, Budapest, Hungary. hal-04407632

HAL Id: hal-04407632

<https://hal.science/hal-04407632>

Submitted on 20 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Relation of a test for assessing pre-service primary teachers' pedagogical content knowledge to other educational variables

Malte Lehmann¹, Bettina Roesken-Winter¹ and Lars Jenßen¹

¹Humboldt-Universität zu Berlin, Germany; malte.lehmann@hu-berlin.de

PCK is a central component of the professional competence of (pre-service) primary teachers in the field of mathematics, and can be differentiated into different facets. In order to examine this knowledge facet in relation to other components of professional competence (e.g., emotions, beliefs, action planning, instructional quality), an economical and freely available test procedure that sufficiently fulfils common quality criteria is needed. This paper presents partly the validation of such a test. According to test standards and in order to prove whether test scores allow for conclusions on pre-service teachers' PCK, relationships to other variables were examined. The results provide validity arguments, indicating a sufficiently high reliability of the test, and support theory-conform conclusions based on the test scores.

Keywords: Pre-service primary teachers, pedagogical content knowledge, test development, validation.

Introduction

Mathematics pedagogical content knowledge (PCK) is an essential facet of teachers' professional knowledge, along with mathematics content knowledge (CK) and general pedagogical knowledge (Shulman, 1986). In larger studies, PCK is usually measured by tests, in studies with smaller samples, it is sometimes measured in qualitative ways (Depaepe et al., 2013). Overall, it can be stated that a freely available test procedure for assessing PCK of pre-service primary teachers is missing. Such an economical, reliable and valid test would allow for investigating connections to other components of professional competence such as beliefs or situation-specific skills.

This paper presents a test for assessing PCK of pre-service primary teachers and describes the relation of the test to other educational variables. First, the construct PCK is theoretically reviewed with regard to its content, and structure. Then the concept of validation and the development of the test are presented. Based on this, the goals and the basic assumptions to be tested in the validation process are presented. Finally, the evidence reported on the basic assumptions is considered and discussed in an integrative manner.

PCK of prospective primary school teachers

The seminal work of Shulman (1986) introduced PCK as a central component of the cognitive dispositions of teachers' professional knowledge. Building on this, a variety of different definitions of PCK emerged, each with a different focus. Depaepe et al. (2013) identified several components of PCK in their systematic review, including knowledge of (mis)conceptions of students, mathematical tasks and cognitive demands, curricular knowledge as well as content knowledge and pedagogical knowledge. Another established conceptualization was contributed by the COACTIV project (Baumert, Kunter, et al., 2010). Here, PCK was divided into subdimensions, encompassing knowledge about the diagnostic potential of tasks, the cognitive demands and implicit knowledge

prerequisites of tasks, their didactic sequencing, and the long-term curricular arrangement of material, knowledge about student conceptions (misconceptions, typical errors, strategies), and diagnosing student knowledge and comprehension processes, as well as methodological knowledge (Baumert & Kunter, 2013). Comparably, Ball et al. (2008) structured the construct in their MKT framework into the subdimensions knowledge of content and students, knowledge of content and teaching, and knowledge of content and curriculum. Another model that considers the work of Ball et al. (2008) but emphasizes more the specialization of teachers' knowledge and puts teachers' beliefs in a central position is the MTSK model (Carrillo-Yañez et al., 2018). In this model, PCK is structured into the three dimensions: knowledge of mathematics teaching, knowledge of features of learning mathematics, and knowledge of mathematics learning standards.

Drawing on these established subdimensions of PCK of pre-service teachers, and in consideration of (national) specifics of teacher education we opted for the following four facets of PCK: (1) *Students' learning and understanding* (e.g., number aspects of natural numbers); (2) *Curricula and contexts* (e.g., general mathematical competencies as described in national educational standards of primary education); (3) *Tasks and materials* (e.g., use of digital media in mathematics education); (4) *Teaching and instruction* (e.g., representations for unitizing and decomposing in the decimal system). In our test development, the goal was not to develop an instrument that can empirically distinguish these four dimensions, but to use the dimensions as a heuristic for constructing test items.

Development of the PCK test

Our aim was to develop an economical, reliable, and valid measures of pre-service teachers' PCK. To meet the demands, we decided to follow a cognitive and rather static perspective on PCK, comparable to the COACTIV project (Baumert, Kunter et al., 2010) or Ball et al. (2008). With the MaTe test (Assessment of *Mathematics* PCK of pre-service primary *Teachers*) we aim at capturing PCK both in the context of larger studies that allow for IRT modeling and in studies with smaller sample sizes that map test scores according to Classical Test Theory. Furthermore, we were interested in developing an economical test with regard to the number of items. All test items were therefore constructed in single-multiple-choice format with one attractor and three distractors (see Figure 1).

Each item also contained the option "I do not know the answer". Checking this option or a distractor was coded 0 for incorrect in scoring, whereas choosing the attractor was coded 1 for correct in scoring. The design process was organized in several phases. First, the authors and colleagues were trained regarding the basics of test construction. Second, based on the framework described above, each person composed items related to his or her area of research or teaching. Thereby, items from former exams, from the literature used for teacher education, or from published items from other research projects served as a basis for the item construction. A total of 110 items were constructed. Third, an internal round of review by the project team was conducted, finally leading to identifying 59 items that satisfied the content and formal requirements of the test construction. Fourth, the item pool of 59 items was examined by experts with regard to the quality of content representation at the item level. Items that could not be certified as having sufficient content representation were eliminated. Finally, the remaining 44 items were analyzed for their empirical quality in a pilot study with $n = 314$

prospective primary school teachers at two universities. Based on the piloting results, 24 items were selected for the test.

<i>You would like to discuss the cardinal number aspect of natural numbers in your class. Which situation is very suitable for explaining this to the students?</i>
Cem runs a race against four friends. He finishes ahead of Tom and Maria, but behind Clara. What is his finishing position? (Distractor 1)
Azra lives two kilometers away from school. Maja lives three kilometers away in the opposite direction. How far is it from Azra to Maja? (Distractor 2)
The post office uses postal codes to deliver items to the correct city. How do these differ? (Distractor 3)
Anna is supposed to buy four apples and five pears for her father. How many items of fruit should she bring? (Attractor)

Figure 1: Example of the MaTe-test. An item of the dimension *Students' learning and understanding*

Relation of the PCK test to other educational variables

The overarching goal of the present study was gathering empirical evidence by a comprehensive validation. In accordance with Kane's (2013) argumentation-based approach to validation, we pursued the following overarching research question:

RQ: Do the MaTe test scores show differential relationships to other variables?

We tested several hypotheses within the evidence category *relationships to other variables*. In order to use the results of the MaTe test in the sense of a proxy for pedagogical content knowledge, it must be shown that there are correlations between the MaTe test score and constructs related to PCK. According to the theoretical conceptualization, it can be assumed that pre-service primary teachers acquire their PCK throughout the course of their university studies and that this acquisition is convergently or divergently related to achievements from their school years. Based on theoretical assumptions (Ball et al. 2008; Shulman 1986), a convergent relationship between CK and PCK can also be postulated. Studies indicated that these facets tend to be highly correlated. For instance, Voss et al. (2011) found CK and PCK of pre-service secondary teachers highly correlated with $r = .91$ and Blömeke, Jenßen et al. (2015) identified for the CK and PCK of prospective early childhood teachers a latent correlation of $r = .73$. In another study, Blömeke et al. (2010) revealed a manifest correlation of $r = .62$ for preservice primary teachers' CK and PCK. The magnitude of this correlation indicated that the two constructs differ not only theoretically but also empirically. Accordingly, we postulate the following hypothesis:

(a) The MaTe test scores are highly positively related to pre-service primary teachers' CK. Considering the evidence for pre-service teachers in Germany and specifically the study results of Blömeke et al. (2010), we assume a correlation not above .9 but rather around .6.

Our next hypotheses refer to the theoretical assumption that cognitive constructs are generally interrelated among learners (Rindermann, 2007). That is, domain-specific variables (e.g., grades in mathematics) should show a correlation with the MaTe test scores, whereas variables of other domains (e.g., grades in German) or variables comprising a variety of domain-specific and domain-distant achievements (e.g., average high school diploma) should show divergent correlations. Given that higher grades represent lower achievement, the following hypotheses can be stated:

(b) The MaTe test scores show positive correlations with the last school grade in mathematics. This assumption reflects the domain specificity of the MaTe test with respect to content. However, since the school grade in mathematics does not reflect any PCK, no strong correlation is assumed.

(c) The MaTe test scores have no to small positive correlations with the last school grade in German. This assumption can be justified by the fact that the last school grade in German can be seen as a proxy for language competence and reading comprehension and that test items on professional knowledge usually contain longer text sections. Since the MaTe test does not include case vignettes or longer situation descriptions, we assume that the correlation may be small at most.

(d) There are no to small positive correlations with the average grade in the *Abitur* (high school diploma in Germany). Studies indicate correlations between knowledge tests and the average *Abitur* grade of student teachers (e.g., Kleickmann et al. 2013). Since the MaTe test is intended to capture knowledge that is not acquired during schooling and the baccalaureate average score nevertheless includes a bundle of cognitive skills, we assume no correlation or at most a small correlation for the two variables.

The program usually lasts six semesters for the bachelor's degree and four semesters for the master's degree. Based on the theoretical conception of professional knowledge, it can be deduced that as the number of semesters at university increases, professional knowledge also increases. Therefore, the following hypothesis can be formulated:

(e) MaTe test scores are positively related to the number of semesters.

Methods

To test the hypotheses, $n = 440$ pre-service primary teachers at one university filled out the MaTe test, the majority (80.1%) of participants were female (19.4% male, 0.5% diverse). Participants were recruited from all semesters of teacher education, but were at least in their 3rd semester (Mean (M) = 4.92, Standard deviation (SD) = 1.64). This is also reflected in the spread of age ($M = 24.40$, $SD = 5.90$). Participants completed the 24 MaTe test items online. The students did not receive any compensation for their efforts and participated in the survey voluntarily.

Since not all participants of the total sample completed all instruments, subsamples were examined to test the different assumptions. Subsample 1 included $n_1 = 223$ pre-service primary teachers and was used to test hypothesis (a). The participants came from one lecture at the end of the bachelor's program and thus had already participated in courses on subject-specific and didactical content of arithmetic, geometry, and stochastics. Subsample 2 consisted of $n_2 = 217$ pre-service primary teachers and was used to test assumptions (b) – (e). The majority of participants were in the bachelor's program (85%), and the minority in the master's program (15%). In terms of gender ratio and age distribution, the two subsamples did not differ descriptively from the overall sample.

All participating pre-service teachers completed all 24 items of the MaTe test. Participants in subsample 1 additionally completed the KomMa-MCK test to assess their CK (Blömeke, Jenßen, et al., 2017). This test was originally intended to assess the content knowledge of early childhood teachers, but it was constructed at the level of primary teachers. Accordingly, the suitability of the test for pre-service primary teachers could be supported in a previous study (Dunekacke et al., 2014).

Also, KomMa-MCK was found to provide test scores that can be empirically distinguished from general cognitive abilities (Jenßen et al., 2019). The test comprises 22 items in a single-multiple-choice and open-response format and measures CK in four different content areas (arithmetic; geometry; stochastics; and quantities, patterns, and structures), considering process-related competences (reasoning; representing; modeling; problem solving; and communicating). Two items were not used because a digital implementation was not possible. In addition to the MaTe test scores and the number of semesters, we collected information about participants' last school grade in mathematics, last school grade in German, and *Abitur* grade for participants in subsample 2. Both the two tests and the collection of the additional variables were conducted online using *LimeSurvey* software. The students had no time constraints.

In order to investigate the hypotheses, participants' ability parameters were correlated with the values of the respective variables. In our analysis of the correlations between the MaTe score and the school grades, the scaling for school grades commonly used in Germany was changed: Instead of "1-very good" to "6-insufficient", we used "1-insufficient" to "6-very good" (*Mplus 8* software, Muthén and Muthén, 2017) for the analyses. The analyses with *Mplus* allowed for accounting missing values according to the Full Information Maximum Likelihood method.

Results

Table 1 shows the descriptive results for the investigated variables. Table 2 shows the correlations of these variables with the MaTe test score. It can be seen that hypothesis (a) (high positive correlation with CK) can be considered confirmed. Likewise, the data supports hypothesis (b) (medium positive correlation with the last school grade in mathematics) and hypothesis (e) (positive correlation with semester number). However, no evidence was found for hypotheses (c) (small positive correlation with the last school grade in German) and (d) (small positive correlation with the *Abitur* grade).

Table 1: Descriptive results

	<i>M</i>	<i>SD</i>
KomMa-MCK-test score	15.93	3.70
Last school grade in mathematics	2.46	1.02
Last school grade in German	2.33	0.83
Average <i>Abitur</i> grade	2.46	1.96

Table 2: Correlations of the variables with the MaTe test score

	KomMa-MCK-test score	Last school grade in mathematics	Last school grade in German	Average <i>Abitur</i> grade	Semester
MaTe-test score	.58 ($p < .01$)	.26 ($p < .01$)	.08 ($p = .09$)	.05 ($p = .46$)	.36 ($p < .01$)

Summary and discussion

As recommended by the AERA, APA, and NCME (2014), in the following we conduct an integrative review of the presented validity evidence. Our goal was to investigate whether the test score of the MaTe test can be interpreted in terms of explanations. Explanation here means that the test scores reflect the construct PCK of pre-service primary teachers.

With regard to *the relationships to other variables*, the test score showed consistent hypothesis-compliant properties. The correlation pattern corresponded to the hypotheses formulated in advance. In the case of the last school grade in German, our data suggested that the completion of the MaTe test items is not associated with linguistic performance on part of the participants. Here, a validation with a test for the measurement of linguistic competencies would be more appropriate in the future in order to substantiate our findings. It is questionable whether the non-significant result with regard to the *Abitur* grade already shows that the MaTe test contains tasks whose content is not acquired in school, and exclusively captured during university teacher education. Adding the found correlation to the number of semesters, this assumption is supported, even when our study design does not allow for falsification of this assumption (Hartig et al., 2020). Longitudinal studies would be needed in the future to strengthen this validity argument.

Regarding the relevance of such a test, we would like to point out that the test can be applied in two ways. First, the MaTe test can be used for evaluating students' development over the course of their studies. The small correlation with the last school grade in mathematics and the correlation with the number of semesters indicates that the test captures knowledge that is acquired during university studies. It can be used to track developments over longer periods of time. Second, as many studies, especially qualitative studies on prospective mathematics teachers' competencies, examine very specific aspects such as diagnostic competence, they do not provide information on teachers' overall PCK. The MaTe test can thus be used as a covariate, or a proxy for participants' PCK and help to paint a comprehensive picture of teachers' competencies.

Finally, we like to conclude by highlighting that a validation process is never complete (Kane, 2013; AERA, APA, and NCME, 2014), as other inferences that have not yet been studied are always possible. From our perspective, the following future validation strategies represent useful extensions of the present study. At the item level, it may be interesting to investigate response processes in the form of cognitive labs to analyze what specific PCK is used when working on the MaTe items (Padilla & Leighton, 2017). Also, we could examine how results on the MaTe test specifically follow periods of practice, as these are specific PCK learning opportunities that are considered to have large effects for knowledge acquisition during undergraduate teacher education programs (Phelps et al., 2019).

References

- AERA, APA & NCME (=American Educational Research Association, American Psychological Association, & National Council on Measurement in Education). (2014). *Standards for educational and psychological testing*. American Psychological Association.
- Ball, D., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, 59(5), 389–407.

<https://doi.org/10.1177/0022487108324554>

- Baumert, J., & Kunter, M. (2013). The COACTIV model of teachers' professional competence. In: Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S., Neubrand, M. (Eds.), *Cognitive Activation in the Mathematics Classroom and Professional Competence of Teachers. Mathematics Teacher Education*, vol 8. Springer, Boston, MA. https://doi.org/10.1007/978-1-4614-5149-5_2
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M., & Tsai, Y. M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, 47(1), 133–180. <https://doi.org/10.3102/0002831209345157>
- Blömeke, S., Gustafsson, J.-E., & Shavelson, R. J. (2015). Beyond dichotomies: Competence viewed as a continuum. *Zeitschrift Für Psychologie*, 223(1), 3–13. <https://doi.org/10.1027/2151-2604/a000194>
- Blömeke, S., Jenßen, L., Grassmann, M., Dunekacke, S., & Wedekind, H. (2017). Process mediates structure: Relation of preschool teacher education and preschool teachers' knowledge. *Journal of Educational Psychology*, 109, 338–354. <https://doi.org/10.1037/edu0000147>
- Blömeke, S., Kaiser, G., Döhrmann, M., Suhl, U., & Lehmann, R. (2010). Mathematisches und mathematikdidaktisches Wissen angehender Primarstufenlehrkräfte im internationalen Vergleich [Mathematical content and mathematical pedagogical content knowledge of pre-service primary teachers in international comparison]. In S. Blömeke, G. Kaiser & R. Lehmann (Eds.), *TEDS-M 2008: Professionelle Kompetenz und Lerngelegenheiten angehender Primarstufenlehrkräfte im internationalen Vergleich* [Professional competence and learning opportunities of pre-service primary teachers in international comparison] (pp. 197–251). Waxmann.
- Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C.-P., & Loeffel, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: An experimental study. *American Educational Research Journal*, 26(4), 499–531. <https://doi.org/10.3102/00028312026004499>
- Carillo-Yañez, J., Climent, N., Montes, M., Contreras, L. C., Flores-Medrano, E., & Escudero-Ávila, D. (2018). The mathematics teacher's specialised knowledge (MTSK) model. *Research in Mathematics Education*, 20(3), 236–253. <https://doi.org/10.1080/14794802.2018.1479981>
- Depaepe, F., Verschaffel, L., & Kelchtermans, G. (2013). Pedagogical content knowledge: A systematic review of the way in which the concept has pervaded mathematics educational research. *Teaching and Teacher Education*, 34, 12–25. <https://doi.org/10.1016/j.tate.2013.03.001>
- Dunekacke, S., Jenßen, L., Eilerts, K., & Grassmann, M. (2014). Operationalisierung des mathematikbezogenen Wissens angehender frühpädagogischer Fachkräfte [Operationalization of mathematics-related knowledge of pre-service early childhood educators]. In Schuler, S., Streit, C., Wittmann, G. (Eds.) *Perspektiven mathematischer Bildung im Übergang vom Kindergarten zur Grundschule* [Perspectives on mathematical education in the transition from kindergarten to primary school]. (pp. 207–222). Springer Spektrum. https://doi.org/10.1007/978-3-658-12950-7_14

- Hartig, J., Frey, A., & Jude, N. (2020). Validität von Testwertinterpretationen [Validity of test value interpretations]. In H. Mossbrugger & A. Kelava (Eds.), *Testtheorie und Fragebogenkonstruktion* [Test theory and questionnaire construction] (pp. 529–545). Springer Nature. https://doi.org/10.1007/978-3-662-61532-4_21
- Jenßen, L., Dunekacke, S., Gustafsson, J.-E., & Blömeke, S. (2019). Intelligence and knowledge: the relationship between preschool teachers' cognitive dispositions in the field of mathematics. *Zeitschrift Für Erziehungswissenschaft*, 22, 1313–1332. <https://doi.org/10.1007/s11618-019-00911-2>
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1–73. <https://doi.org/10.1111/jedm.12000>
- Kleickmann, T., Richter, D., Kunter, M., Elsner, J., Besser, M., Krauss, S., & Baumer, J. (2013). Teachers' content knowledge and pedagogical content knowledge: The role of structural differences in teacher education. *Journal of Teacher Education*, 64(1), 90–106. <https://doi.org/10.1177/0022487112460398>
- Krathwohl, D. R. (2002). A revision of Bloom's Taxonomy: An overview. *Theory and Practice*, 41(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2
- Krauss, S., Baumert, J., & Blum, W. (2008). Secondary mathematics teachers' pedagogical content knowledge and content knowledge: Validation of the COACTIV constructs. *ZDM - Mathematics Education*, 40(5), 873–892. <https://doi.org/10.1007/s11858-008-0141-9>
- Muthén, L., & Muthén, B. (2017). *Mplus user's guide* (8th Ed.). Muthén & Muthén.
- Padilla, J.-L., & Leighton, J. P. (2017). Cognitive interviewing and think aloud methods. In B. D. Zumbo & A. M. Hubley (Eds.), *Understanding and investigating response processes in validation research* (pp. 211–228). Springer International Publishing. https://doi.org/10.1007/978-3-319-56129-5_12
- Phelps, G., Howell, H., & Liu, S. (2019). Exploring differences in mathematical knowledge for teaching for prospective and practicing teachers. *ZDM - Mathematics Education*, 2008. <https://doi.org/10.1007/s11858-019-01097-x>
- Rindermann, H. (2007). The g-factor of international cognitive ability comparisons: The homogeneity of results in PISA, TIMSS, PIRLS and IQ-tests across nations. *European Journal of Personality*, 21, 667–706. <https://doi.org/10.1002/per.634>
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14. <https://doi.org/10.3102/0013189X015002004>
- Voss, T., Kunter, M., & Baumert, J. (2011). Assessing teacher candidates' general pedagogical/psychological knowledge: Test construction and validation. *Journal of Educational Psychology*, 103(4), 952–969. <https://doi.org/10.1037/a0025125>