

Structure of pedagogical content knowledge in maths teacher education

*Pascal Kilian^{1,2,3}, Judith Glaesser^{2,3}, Frank Loose^{3,4}
and Augustin Kelava^{2,3}*

Abstract

A key challenge in maths education research is the identification of content knowledge (CK) and pedagogical content knowledge (PCK). Our study uses German pre-service teachers to investigate these dimensions and to differentiate PCK further into the domains of instructional and diagnostic competence. Empirical results support the existence of these two domains and show that they can be found in a very content related context. A bifactor model is used to illustrate the within-structure of PCK. The model's validity is discussed referring to different types of students and to relevant validity coefficients of scales of other studies. We discuss implications for the theoretical foundation of the organization of teacher training.

Keywords: mathematics, PCK domains, pedagogical content knowledge, maths teacher education

¹ Correspondence concerning this article should be addressed to: Pascal Kilian, Eberhard Karls Universität Tübingen, Methods Center, Hölderlinstr. 29, 72074 Tübingen, Germany.
E-Mail: pascal.kilian@uni-tuebingen.de

² Methods Center, University of Tübingen

³ Tübingen School of Education

⁴ Department of Mathematics, University of Tübingen

Introduction

In maths teacher education, students face potential problems experienced at the transition from the way mathematical contents are taught at university to how they are taught in school (second part of "double discontinuity" in Klein (1908), Winsløw & Grønbaek (2014), or "expert blind spot hypothesis" in Nathan and Petrosino (2003)).

This study addresses this problem by investigating the structure of pedagogical content knowledge as something which might be suitable for supplementing the content knowledge education at university.

Shulman (1986, 1987) suggested a general knowledge base for teachers. The distinction between general pedagogical knowledge (GPK), content knowledge (CK) and pedagogical content knowledge (PCK) has become accepted in the literature (e.g., Baumert & Kunter, 2006). One of the aspects covered by PCK is to comprise the ability of transferring abstract expertise to the sort of knowledge which is relevant for teaching. PCK covers a range of areas from content-related elements of teacher knowledge (referred to as the content-related part of PCK) to those elements relating to students and classroom which may be understood as the pedagogical end of the spectrum. Despite its popularity, the wide range of the term – as used by Shulman – lacks definition and empirical investigation (Ball, Thames, & Phelps, 2008; Depaepe, Verschaffel, and Kelchtermans, 2013) and indicates a multidimensional construct and it therefore seems appropriate to use a model that takes this dimensionality into account.

The question arises to what extent (content-related) PCK includes special mathematical knowledge which should be taught in the context of PCK and in addition to the important subject-matter knowledge lectures (e.g. Ball, Lubienski, & Mewborn, 2001) of the study program (and thus – in the curriculum of the study program – separated of the CK education).

The concentration on its content-related tasks is a new approach of investigating PCK, as this aspect formed only a small part of the investigation of a general picture in earlier studies (e.g. N. Buchholtz, Kaiser, & Stancel-Piatk, 2011), or it was assigned to CK (e.g. Ball et al., 2008). The domains of PCK in particular have hardly been studied.

Related work

During the last decades, several studies on the relation of CK and PCK were conducted to provide some of the empirical evidence which was considered to be lacking by Baumert and Kunter (2006). Those studies measure CK and PCK separately or their relation to each other among other aspects of the broad professional competence of teachers of mathematics (e.g. COACTIV, Kunter, Baumert, Blum, Klusmann, Krauss, & Neubrand (2011), Kunter & Baumert (2011); TEDS-M, Blömeke, Hsieh,

Kaiser, and Schmidt (2014); MT21, Blömeke (2011); Blömeke, Kaiser, and Lehmann (2008); Krauss, Brunner, et al. (2008)). They are concerned with teachers in service or students at the end of education and focused on contents for lower grades (e.g. Blömeke & Kaiser, 2014; C. Buchholtz, Doll, Stancel-Piatak, Blömeke, Lehmann, & Schwippert, 2011; N. Buchholtz, Scheiner, Döhrmann, Suhl, Kaiser, & Blömeke, 2012).

In those studies, the relationship between CK and PCK was measured using only a single scale for PCK and domains were merely formulated for the construction of tests but have not been empirically verified (e.g. N. Buchholtz et al., 2011). The question about the dimensionality of PCK remains open, but this point is of special interest at the beginning of teacher education because it may be important to know which parts develop together and should be taught together.

Besides other reconceptualization of teachers' PCK (e.g. Cochran, DeRuiter, & King, 1993; Grossman, 1990; Marks, 1990), the framework of *mathematical content knowledge for teaching* (MKT) (e.g. Ball et al., 2008; Hill, Ball, & Schilling, 2008) which covers both CK and PCK, also reveals interesting insights. Ball et al. (2008) define MKT as "the mathematical knowledge needed to carry out the work of teaching mathematics" (Ball et al., 2008, p. 395).

For the component that includes content-specific knowledge, the framework of Schoenfeld and Kilpatrick (2008) was used. This framework is influenced by Shulman (1987) and distinguishes between *mathematics content knowledge* (MCK) and *mathematics pedagogical content knowledge* (MPCK).

As our focus is on maths specific CK and PCK, we use the abbreviations MCK and MPCK.

The present study

As part of our *Maths Teacher Education Study* (MatTES), this study intends to investigate the inner structure of MPCK and how it can support teacher education in Germany. This approach differs from that taken in other studies (e.g. Blömeke et al., 2011; Blömeke, Kaiser, & Lehmann, 2010; N. Buchholtz et al., 2012, Baumert & Kunter, 2011; Krauss, Neubrand, Blum, & Baumert, 2008) in that it does not intend to measure correlations of MPCK and MCK in general. Instead, the aim of the study is to investigate dimensions within MPCK and to compare latent mean scores of different groups differentiated by experience and degree program with the aim of validating the test.

The study design of MatTES focusses on two particular aspects. The first is a concentration on the teacher education for the academic track (upper secondary education) where the type of content studied is close to academic mathematics

The second aspect concerns the stage of training. In Germany, teacher pre-service education takes place in two phases. Our focus is on the beginning of the first,

university-based phase. It involves formal education in MCK, MPCK and GPK as well as internships in school in two subjects.

With regard to MCK, our study agrees with the four stages of mathematical understanding in Baumert and Kunter (2011). A profound understanding of the mathematical background of the subject matter taught in school is seen as a prerequisite for MPCK within MCK. Our study refers to this level of abstraction as *school-relevant mathematical content knowledge* (schoolMCK) in contrast to *academic mathematical content knowledge* (academicMCK). For simplification, readers may think of schoolMCK as the CK described by Shulman (or CCK in Ball et al. (2008)).

MPCK is coming into play in the content-related interaction between the school teacher and her students. We therefore expect to see two domains of MPCK which relate to the two directions of this interaction. Those domains are based on the original key components – knowledge of instructional strategies and representations (*instruction*) and knowledge of students' (mis)conceptions (*diagnostic competence*) – in Shulman (1986) (see also KCT and KCS in Ball et al. (2008)).

The domain of *instruction* refers to the teacher's ability to prepare and communicate. This process includes the integration of the school content into an academic context as well as the extraction and presentation of subject matter (see *tasks* in Baumert and Kunter (2011)) in a form suitable to a specific group of students (depending on the grade and the previous knowledge of those specific students). Instruction includes knowledge about typical misunderstandings, ways to avoid them or resolve them if they arise. Note that we include misunderstandings here (see Shulman's knowledge of students' (mis)conceptions). We include this notion here because it is important for the selection of strategies to understand where problems might occur. In these domains we also include content-related parts (also described in SCK, Ball et al. (2008)). For example, the mathematical knowledge required for the tasks of "finding an example to make a specific mathematical point" and "modifying tasks to be either easier or harder" (Ball et al., 2008, p. 400) can be seen as the foundation or the content-related part of the *instruction* domain and thus it is included in our framework. This domain can be compared to Shulman's knowledge of instructional strategies/representations

The process which is concerned with the domain *diagnostic competence* includes the identification of underlying sources of mistakes and the evaluation of individual states of knowledge from the responses of the students (described as a component of SCK by Ball et al. (2008)). Again, we include this content-related part in our domain within MPCK. In this context, a response can be any kind of feedback the teacher receives from his or her students, for example in oral discussions in the classroom or through test results. In that sense *diagnostic competence* can be seen as part of the more general diagnostic competence (described for example in Ohle and McElvany (2015)). Close to Shulman's PCK the general diagnostic competence can be defined as the ability to judge students' performance level correctly as well as the correct estimation of the difficulty of tasks and materials (see e.g. McElvany, Schroeder, Baumert, Schnotz, Horz, & Ullrich, 2012). Included in that general framework, *diagnostic competence* qualifies as the part that is concerned solely with the judgement of students' responses

to specific contents and tasks. Diagnostic competence is often seen as an important component of teachers' competencies alongside PCK. However due to the closeness of our notion of *diagnostic competence* to specific contents it is seen as part of MPCK. This domain can be compared to Shulman's knowledge of students' (mis)conceptions.

Even though those two domains may be regarded as different aspects of teachers' knowledge, in practice they are not completely separate. The domains are connected through the teacher's process of reflection on his or her own actions.

Those domains were discussed theoretically with experts both on MCK (math lecturers at university) and MPCK (lecturers at teacher education institutes).

Research Question and Hypothesis.

As subject matter is often seen as a prerequisite (e.g. Kunter et al., 2011) for MPCK, the question arises whether it is possible to identify domains of MPCK during this phase dominated by subject matter, which exist in addition to MCK. In order to answer this question, we tested two models. Model 1 (see Figure 1a) describes the MPCK domains in addition to the prerequisite of a general MCK domain (here schoolMCK). Model 2 (see Figure 1b) suggests that schoolMCK alone explains the outcomes of the test.

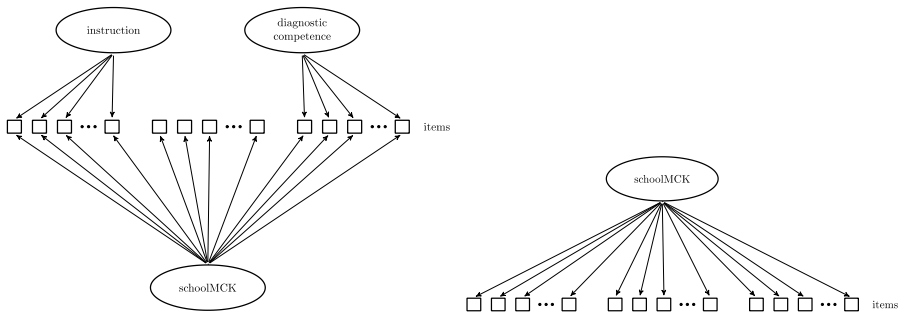


Figure 1.

- a (left) Model with schoolMCK (school relevant mathematics content knowledge) and two domains of MPCK (mathematics pedagogical content knowledge) (model 1),
 b (right) One-dimensional model with schoolMCK (school relevant mathematics content knowledge) as single latent variable (model 2)

In this context, these domains may be seen as a specific mathematical understanding that "goes beyond knowledge of subject matter per se to the dimension of subject matter knowledge for teaching" (Shulman, 1986, p. 9).

Ball et al. (2008) mention the importance of evidence for a possible multidimensionality of mathematical knowledge for the organization of teacher education. "Professional education could be organized to help teachers learn the range of knowledge and skill they need in focused ways" (p. 399). If the domains can be identified, teacher education can be structured accordingly regarding the tasks of MPCK lectures and seminars. While student and classroom related topics within MPCK may only play a minor role in this phase, the content-related parts of the domains should be supported in parallel to the mathematical lectures focusing on MCK.

A summary of the tasks of MPCK lectures in the first phase of math teacher education is shown in Figure 2. One task is to help students extract and adapt the academic maths contents they learn in the lectures in a way which then turns it into the schoolMCK domain. This means that an MPCK lecture should support students' transformation of abstract knowledge into knowledge which is useful for teaching in school. The second task is to produce the mathematical understanding necessary for the domains *instruction* and *diagnostic competence*.

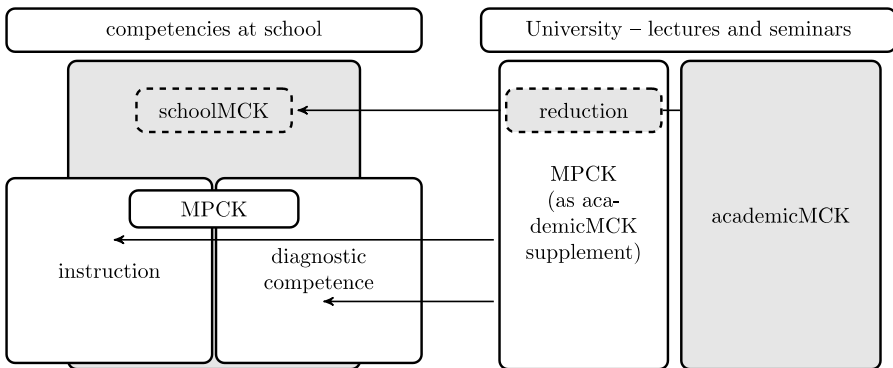


Figure 2.

Tasks of MPCK (mathematics pedagogical content knowledge) lectures at the early stage in university as support of the subject matter education. MPCK as competence in school composed of the two domains instruction and diagnostic competence, schoolMCK (teachers' school-relevant mathematical content knowledge) as underlying requirement.

Hypothesis: Domains of MPCK explain outcomes in addition to MCK in a content-related MPCK test. This means that the proposed model 1 (see Figure 1a) fits the data and the domains explain variance even for controlled schoolMCK, and model fit improves compared to a one-dimensional model (see Figure 1b) with schoolMCK as single latent variable.

Methods

Study design and sample

The study was conducted within the project MatTES which, started at the end of 2014 at the University of Tübingen. For the research question it was necessary to analyze students with different levels of experience in MCK. We focused solely on MCK in the field of calculus (analysis). A broad calculus training is covered by the lectures Analysis 1 through to Analysis 4, usually attended in the first semester through to the fourth semester. A rough classification would be: One-dimensional integral and differential calculus in Analysis 1, multidimensional differential calculus in Analysis 2, multidimensional integral calculus in Analysis 3 and complex analysis (in one variable) in Analysis 4. In the maths major course (B.Sc.), it is compulsory to attend all analysis lectures. The state examination students – pre-service teachers – (state examination is the German graduate degree for teachers) have to attend Analysis 1, 2 and 4, Analysis 3 is not compulsory for this group of students.

For the teacher candidates, two courses in maths specific didactics are required (Didactics 1 and Didactics 2). Students starting in winter semester 2014/15 were free to choose at which point during their studies they attended a program of didactics lectures. Students starting in winter semester 2015/16 were scheduled to take didactics in their second year (third semester), at which point they had already acquired basic knowledge of academic maths. Therefore, the participants of the Analysis 2 lecture had not yet attended a didactics lecture at the start of our study. At the end of winter semester 2015/16, the Didactics 2 course was taught as a one-week intensive seminar.

The empirical study was conducted drawing on participants of two lecture courses and one seminar course. For the main sample, we analyzed students of the two lecture courses Analysis 2 and Analysis 4 (compulsory for teacher candidates) in the summer semester of 2016. These lectures are attended both by students taking either maths or physics as their major subject (in the following we will refer to those two groups together as "B.Sc.") and by teacher candidates (referred to as "Teacher"). Nevertheless, this is a homogeneous sample regarding the study program because their maths components do not differ at this stage and it is a homogeneous sample within the preservice teachers because until this stage the students mostly focus on the maths education with hardly any additional education in didactics. In addition, students attending the seminar for maths specific didactics (Didactics 2) were included (referred to as

"Tsem"). The participants were teacher candidates in the fourth semester and above. Some of them had already completed an internship at school.¹ The sample contained $n = 256$ students, 112 of which were recruited from the Analysis 2 lecture and 116 from Analysis 4. The sample from Didactics 2 contained 28 students. In addition to the programs of study described here (Teacher and B.Sc.), a few students from other programs attended these lectures, which is why the numbers for Teacher and B.Sc. do not add up to the total. Table 1 shows descriptive statistics. Covariates (e.g., age, school grades) were measured using a questionnaire.

Table 1
Descriptive data for the samples and subsamples

event subsam- ple	Analysis 2			Analysis 4			Didactics 2 Tsem ($n = 28$)	total sample ($n = 256$)
	total ($n = 112$)	Teacher ($n = 63$)	B.Sc. ($n = 31$)	total ($n = 116$)	Teacher ($n = 50$)	B.Sc. ($n = 59$)		
	$M (SD)$	$M (SD)$	$M (SD)$	$M (SD)$	$M (SD)$	$M (SD)$	$M (SD)$	$M (SD)$
Age ^a	21.37 (4.65) ($n = 100$)	20.90 (1.64) ($n = 63$)	22.32 (8.0) ($n = 31$)	22.29 (2.34) ($n = 111$)	22.90 (1.92) ($n = 50$)	21.86 (2.59) ($n = 58$)	24.12 (1.35) ($n = 28$)	22.12 (3.53) ($n = 239$)
High school GPA ^b	1.82 (0.55) ($n = 98$)	1.86 (0.55) ($n = 63$)	1.73 (0.56) ($n = 29$)	1.87 (0.56) ($n = 112$)	1.81 (0.56) ($n = 50$)	1.92 (0.56) ($n = 58$)	1.89 (0.54) ($n = 28$)	1.85 (0.55) ($n = 238$)
High school maths score ^c	12.34 (2.41) ($n = 97$)	12.19 (2.35) ($n = 62$)	12.90 (2.21) ($n = 29$)	12.64 (2.53) ($n = 108$)	12.53 (2.02) ($n = 49$)	12.91 (2.44) ($n = 55$)	12.20 (2.06) ($n = 25$)	12.47 (2.43) ($n = 230$)
Gender (male)	48 % ($n = 100$)	37.1 % ($n = 63$)	64.52 % ($n = 31$)	59.29 % ($n = 113$)	36 % ($n = 50$)	76.27 % ($n = 59$)	32.14 % ($n = 28$)	51.15 % ($n = 241$)

Note: ^a rescaled measure (2016 – (year of birth)). ^b German grade point average (GPA), scores range from 1 to 6 with 1 as the best score. ^c maths score of the final secondary school examination. Scores range from 1 to 15 with 15 as the maximum score.

¹ Even though this does not seem homogeneous all the analysis was executed and the structure was tested without the inclusion of the seminar resulting in no significant differences.

Teacher = teacher candidates; B.Sc. = Bachelor of Science (math and physics); Tsem = math specific didactics students.

Instrument

The competence test employed here was developed for trainee teachers at the beginning of their second year at university and above. Basic knowledge of university maths is required. The test is intended to measure knowledge in MPCK that is close to purely mathematical understanding. It is not intended to measure general MPCK skills in teaching and classroom management, because these skills are expected to be developed during later stages of teacher education. Due to this and the general area of application of MatTES it was necessary to develop a new instrument. We considered employing instruments which had been validated in other studies (e.g., TEDS-M, TEDS-sM). However, these would not have been suitable for our purposes, as the target populations differ. Our participants' state of knowledge precluded the use of technical terms in the field of MPCK. Nevertheless, the development of the items drew on those existing instruments and the best subset was selected.

Different experts were involved in the development of the test. On the mathematics side, the group consisted of lecturers involved in the Analysis lectures. For the MPCK content, our group worked together with colleagues at the teacher education institute Tübingen, which is responsible for the second phase of maths teacher education and which provides the MPCK lectures and seminars at the university.

In the interest of economy, a fairly short test was constructed containing 19 items, all of which are in binary multiple-choice format. More items – including existing items – were tested during the development phase. Due to limited testing time, the resulting instrument contains a selection of the best items. The study was conducted in close connection with the Analysis lectures, therefore the mathematical background for the items is the content of the Analysis introduction lecture program taken in the first semester. Example tasks are presented in the appendix.

Instruction sub-scale.

The instruction sub-scale of the test consists of 5 items. The assignment of those items to this domain was reviewed by experts as well as by pilot testing. In this domain, the test-takers were asked to consider representations of mathematical theories that are suitable for students in school. Reduction of abstract theories as well as knowledge about the learning process and the individual knowledge state of the students are important. In developing the instrument to measure this kind of knowledge we asked questions, for example, about suitable representations of mathematical theories in school.

Diagnostic competence sub-scale.

The 5 diagnostic competence items cover the classification of students' responses into taught contents, identification of possible sources of errors, and interpretation of learning processes. Here again, the assignment was reviewed beforehand. For instance, wrong student responses are presented, and the task is to identify the type or source of the underlying error or misunderstanding. This sub-scale describes a mental representation of dysfunctional cognition which then serves as a starting point for further instruction. For example, test-takers were asked to identify the error underlying a wrong student response by deciding, which of the tasks presented require the same mathematical understanding as the original one. For this task, the same mistake that led to the original wrong response may occur again.

schoolMCK sub-scale.

9 items were developed specifically to measure schoolMCK. These cover content knowledge within the field of analysis in a form that is relevant in school. This dimension is seen as a prerequisite for teaching mathematics and for the MPCK domains. The overall rationale behind these items is that we wished to capture the way participants coped with the problem of balancing the reduction of mathematical contents to a form suitable for teaching in schools without losing accuracy. This can be addressed, for instance by rating unconventional student solutions. For these tasks, the test-takers have to be confident enough in their mathematical knowledge to identify correct answers even if these are written in unconventional language or hidden in unconventional thoughts. They also have to be able to isolate main ideas of mathematical theories and objects to discuss them with the students without going into detail of the underlying academic theory.

Validation.

A collection of TEDS-sM items (N. Buchholtz et al, 2012) was used as comparison items (with a reliability of .72 according to N. Buchholtz et al (2012)). Due to time restrictions, it was not possible to use the full TEDS-sM scale. Nine items within three different tasks were chosen, all from the cognitive dimension *evaluate and create*. The tasks are DS29 and DBJ4 from the topic of *subject matter didactics* and SUG2 from *education didactics*. This choice was made in order to achieve a good fit with the topics of the lectures. Very minor language changes had to be made to ensure that all technical terms were known to our students.

Procedure

The students voluntarily filled in a questionnaire in the second week of the summer semester 2016. The survey was conducted in the tutorials accompanying the lectures which are attended by about 15 to 20 students each. The test time was around 35 min altogether.

Analysis

In the analysis, we used structural equation modeling to fit latent models. First, the bifactor model 1 (see e.g. Reise, 2012) (see Figure 1a), then the alternative model 2 (see Figure 1b) was fitted. Data were analyzed using the lavaan-package (Rosseel, 2012) and robust maximum likelihood estimation was used.

Assessment of model fit.

We tested the fit of the bifactor model 1 and compared it to the unidimensional model 2 to investigate the hypothesis that the two domains of MPCK can explain additional variance in the data. We used the Tucker-Lewis index (TLI), the comparative fit index (CFI) and the root mean square error of approximation (RMSEA) to evaluate goodness of fit (Themessl-Huber, 2014). Values of TLI and CFI greater than .90 and .95 are usually taken to indicate acceptable and excellent fits to the data, respectively. RMSEA values smaller than .60 would show a reasonable fit. In addition, we report the c^2 (Chi-square) test statistic, the ratio of the c^2 deviance and the degrees of freedom. $c^2/df < 2$ values are regarded as good fit.

To compare the nested models, we followed the suggestions of Chen (2007) that a model should be favored if incremental fit indices such as the CFI increases by more than .015 compared to the more parsimonious model. In addition, we executed a c^2 -difference test. Although measurement invariance (Meredith, 1993) should be ensured to conduct comparisons between the sub-groups (semesters), this was not possible with our data due to the small numbers of participants in the subgroups.

Comparison of latent means.

Means on the latent variables were calculated for different majors, different cohorts and gender. We expected to see no differences for gender and no differences between the subgroups of the second semester cohort. At this stage, none of the students had attended didactic lectures and the teacher candidates have not yet had any practical experience. Differences may be expected to occur within the fourth semester cohort. Here, the teacher candidates attending the Analysis 4 lecture may have attended a basic didactics course. The students taking the didactics seminar are more experienced

in didactics as some of them have already completed the practical phase at school. Therefore, we expect to see differences between B.Sc. students with majors in maths or physics and teacher candidates who are either taking Analysis 4 or the didactics seminar.

Validity coefficients.

The predicted values for the latent variables were used to compute correlations with other scales and school grades. For the scaling of the TEDS-sM anchor items a one-dimensional model was fitted using lavaan (Rosseel, 2012), and the person parameters were predicted.

Results

In this section we present the results for our two models. Model 1 (Figure 1a) is a bifactor model (see e.g. Reise, 2012). The alternative model 2 (Figure 1b) is a unidimensional g-factor model. We calculated these models to examine whether the instruction and diagnostic competence sub-scales explain additional variance in the data compared to the g-factor model. Therefore, results of the comparison are also presented. As measure for the reliability of the factors schoolMCK, instruction and diagnostic competence Cronbach's α was used. The values were good (.8) and acceptable (.72 and .7).

Comparison of two alternative models

The bifactor model 1 (see Figure 1a) has a acceptable fit to the data, with respect to the relative fit indices (CFI = .935, TLI = .921), absolute fit indices (RMSEA = .023) as well as $c^2/df = 1.133$. Table 2 presents fit results for both models.

Table 2
Model fit statistics for both models

Model	CFI	TLI	RMSEA	c^2	df	c^2/df
bifactor model 1	.935	.921	.023	158.55	140	1.133
g-factor model 2	.799	.771	.039	207.766	150	1.39

Note: CFI = comparative fit index, TLI = Tucker-Lewis index, RMSEA = root mean square error of approximation

The fit of the alternative model 2 (see Figure 1b) is inferior to that of model 1 (see Table 2) according to the relative fit indices (CFI = .799, TLI = .771) and the absolute fit indices (RMSEA = .039) and $c^2/df = 1.39$. The CFI difference is $\Delta CFI = 0.136$ (Robust) which favors model 1 as it is greater than .015 (Chen, 2007). The results of the scaled c^2 difference test (see Rosseel, 2012; Satorra, 2000) are presented in Table 3. The test is highly significant for model 1. The wording and content of some items, within one task, was similar. Thus, in both models two residual covariances within the diagnostic competence domain had to be calculated. Each residual covariance occurs between two items within one task (see appendix Figures A2 and A3). Estimated covariances for items in Figure A2 were .41 (model 1) and .52 (model 2). Those in Figure A3 were .54 (model 1) and .55 (model 2).

Table 3
Scaled c^2 Differences Test

Model	df	c^2	$\Delta\chi^2$	Δdf	$Pr(> c^2)$
model 1	140	132.29			
model 2	150	191.81	44.6	7.9908	4.344*10 ⁻⁷

Comparison of means

Comparisons of latent dimension means for different subgroups were tested. The results are presented in Table 4 including effect sizes and results of one-tailed t -tests.² The teacher candidates performed significantly worse in the area of school-relevant content knowledge (t -value = 2.0, df = 52.76, d = 0.46), but significantly better in the instruction sub-scale (t -value = -2.59, df = 75.28, d = 0.52). On the instruction sub-scale, the teacher candidates in the didactics seminar also outperformed the less experienced teacher candidates in the Analysis 4 lecture (t -value = -1.83, df = 72.86, d = 0.40). As expected, we found no differences between different majors in the second semester. Note that there were no differences between teacher candidates and students with maths as their major subject at this early stage in their performance on the subject matter part, the schoolMCK sub-scale. On the other hand, for the fourth semester students, significant differences were found between the students with maths or physics as their major subjects and the more experienced teacher candidates of the didactics seminar.

² Possible gender differences were examined, but none of the results were significant.

Table 4
Comparison of means table with t statistics

2. semester					
	Teacher (n = 63)	B.Sc. (n = 31)			
	<i>M (SD)</i>	<i>M (SD)</i>	<i>t-value (df)</i>	<i>p-value</i>	<i>d</i>
sMCK	-0.05 (0.66)	0.03 (0.48)	-0.71 (79.4)	0.24	0.13
I	-0.05 (0.55)	-0.11 (0.58)	0.50 (56.63)	0.31	-0.11
D	-0.13 (0.63)	-0.18 (0.61)	0.34 (60.80)	0.37	-0.08
4. semester					
	Teacher (n = 50)	B.Sc. (n = 59)			
	<i>M (SD)</i>	<i>M (SD)</i>	<i>t-value (df)</i>	<i>p-value</i>	<i>d</i>
sMCK	-0.09 (0.81)	0.03 (0.71)	-0.81 (98.51)	0.21	1.16
I	-0.03 (0.56)	-0.10 (0.57)	0.66 (104.78)	0.25	-0.12
D	-0.02 (0.55)	-0.06 (0.55)	0.46 (104.28)	0.34	-0.07
	B.Sc. (n = 59)	Tsem (n = 28)			
	<i>M (SD)</i>	<i>M (SD)</i>	<i>t-value (df)</i>	<i>p-value</i>	<i>d</i>
sMCK	0.03 (0.71)	-0.30 (0.72)	2.0 (52.76)	0.025	-0.46
I	-0.10 (0.57)	0.17 (0.38)	-2.59 (75.28)	0.01	0.52
D	-0.06 (0.55)	0.04 (0.65)	-0.72 (46.13)	0.24	0.17
	Teacher (A4) (n = 50)	Tsem (n=28)			
	<i>M (SD)</i>	<i>M (SD)</i>	<i>t-value (df)</i>	<i>p-value</i>	<i>d</i>
sMCK	-0.09 (0.81)	-0.30 (0.72)	1.18 (61.89)	0.12	-0.27
I	-0.03 (0.56)	0.17 (0.38)	-1.83 (72.86)	0.04	0.40
D	-0.02 (0.55)	0.04 (0.65)	-0.60 (48.6)	0.35	0.10

Note: sMCK = schoolMCK, I = instruction, D = diagnostic competence, Teacher = teacher candidates, B.Sc. = major in maths or physics, Tsem = participants of the didactics seminar
bold: $p < .05$. Differences occur within the fourth semester groups. At this stage the curricula start to differ (some teacher candidates attended lectures in MPCK).

Validity coefficients

Correlation coefficients (and squared correlation) were calculated between the latent variables of model 1 (sMCKscore, Iscore, Dscore), the scaled TEDS-shortM items (TEDSscore) and the individual's GPA (German grade point average) and maths score (of the final secondary school examination). The results are presented in Table 5. The similarity of the squared coefficients of sMCKscore ($r^2 = .13$ for the maths score and $r^2 = .12$ for the GPA) and TEDSscore ($r^2 = .07$ for the maths score and $r^2 = .07$ for the GPA) with the school variables and their own squared correlation of $r^2 = .14$ indicate the measurement of similar constructs.

Table 5
Table of correlations

	TEDSscore	sMCKscore	Iscore	Dscore	maths score	GPA ^a
TEDSscore	1.00 (1.00)					
sMCKscore	0.37 (0.14)	1.00 (1.00)				
Iscore	-0.06 (0.00)	0.18 (0.03)	1.00 (1.00)			
Dscore	0.15 (0.02)	0.12 (0.01)	-0.10 (0.01)	1.00 (1.00)		
maths score	0.27 (0.07)	0.36 (0.13)	0.03 (0.00)	0.05 (0.00)	1.00 (1.00)	
GPA ^a	0.26 (0.07)	0.35 (0.12)	0.03 (0.00)	0.04 (0.00)	0.63 (0.40)	1.00 (1.00)

Note: ^a inverted, bold: $p < .05$ (two tailed), correlation r and squared correlation (r^2)

Discussion

In MatTES, MPCK was conceptualized with two inner domains – instruction and diagnostic competence – which characterize teachers’ interactions with students. The starting point was a content-dominated approach to MPCK at the beginning of teacher education, where the emphasis is on the development of MCK. For this reason, such a content-related approach is more appropriate than more general conceptualizations of MPCK, which might also include general pedagogical aspects. The approach employed here classifies MPCK as a kind of mathematical understanding that teachers need for interacting with students and it is seen as something different from the mathematical understanding that students gain when attending lectures in pure mathematics. It involves an understanding that provides the ability to transfer abstract knowledge into a form suitable for teaching.

The validity of the instrument was ensured in various ways. First, experts from the teacher education institute in Tübingen and instructors involved in the Analysis lectures worked together during test construction to ensure content validity. Second, the test results show substantial correlations with the items employed in TEDS-sM and with school grades such as the GPA and the final high school maths score. Note that low correlations of Dscore and Iscore (diagnostic competence and instruction) with other scales are due to the fact that we controlled for schoolMCK, and those scores can be seen as any effect remaining after controlling for schoolMCK. Third, the mean differences on the scales conform to expectations based on students’ study programs. At the very beginning of the training there are no differences in competencies. Later in the training, differences in performance were found between students, depending on their program of study, on the schoolMCK scale, and on the instruction scale – the

more experienced teacher candidates perform better than students attending the Analysis lecture.

The identification of MPCK domains in addition to strictly mathematical knowledge within this content-related point of view was carried out by comparing two models – a bifactor model (model 1) with the MPCK domains added to a general dimension of school-relevant mathematical content knowledge (schoolMCK) and the unidimensional model 2 (see Figure 1a and Figure 1b). The results of the model fit analysis favored the bifactor model (model 1), which supports the existence of two MPCK domains in addition to the general schoolMCK domain. The latter is seen as a prerequisite for MPCK. Although the identification of MPCK in addition to MCK had already been analyzed in previous studies, our results are remarkable, because the identification of the different MPCK domains and MCK was undertaken in a highly content-related framework. Within this framework the different constructs can be seen as different kinds of mathematical understanding. Compared to former studies, our results show not only that MPCK and MCK can be distinguished analytically in mathematics teachers, but also that a separation of MPCK and MCK as kinds of mathematical understanding and the different development of these domains in the initial stages of the training, depending on the study program, is supported by the evidence. By separation in this context we do not mean a theoretical separation in the sense of independent dimensions but in the sense of a statistical identification of distinct domains. This shows that MPCK and MCK can be identified separately in the sense of Shulman's PCK as "subject matter knowledge for teaching" (Shulman, 1986, p. 9). This goes beyond the separation of MCK and a comprehensive MPCK dimension – including not only content-related parts but also the general pedagogical point of view – at a later point in the training (as employed by the TEDS-group and COACTIV). In other contexts, a separation seems more obvious because aspects of MPCK which differ greatly from MCK are included. For the investigations in this content-related context, we deliberately scheduled our study at this early stage of the training for two reasons. The first is that at that stage, the focus is on developing subject matter knowledge. The second is that study programs for all students in mathematics are similar at that stage, and in particular there are no lectures and seminars in the MPCK context. The identification of the domains here can then be seen as evidence for the importance of supplemental lectures addressing this kind of mathematical understanding already at this early stage.

The results show no differences between teacher candidates and other students at the beginning of the training. Neither in the MCK nor in the MPCK domains did major differences arise. This was expected due to the homogeneous sample and to their having an identical curriculum in the second semester. In the fourth semester the curricula of the teacher candidates and the other students are different. While students who major in maths concentrate mostly on mathematical lectures, the teacher candidates attend additional lectures and seminars and some of the teachers' sample in the fourth semester had already completed the practical phase in school. The results reflect this specialization in differences of the latent variable means. The more experienced teacher candidates of the didactics seminar outperformed teacher candidates and the

other students in the Analysis 4 lecture on the instruction scale. This could have been caused by the experience those students gained in the practical phase at school, the expert monitoring included in this practical phase, as well as by the didactic seminars and lectures. By contrast, the students of the didactics seminar did less well on the MCK domain – schoolMCK – than the students in Analysis who major in mathematics or physics. This is not surprising in view of the differing curricula in later phases of the training. Depending on the study program the development of competence differs. While the programs of the major in mathematics and physics focus on the subject matter, the focus in the study programs for teacher candidates shifts to a parallel development of knowledge in both MCK and MPCK.

The results may be seen as evidence that it is possible to separate different kinds of mathematical understanding – referred to as MPCK and MCK – as early as the beginning of the training. This result also supports the notion of a mathematical knowledge and skill unique to teaching (described as SCK by Ball et al. (2008)), which should be developed in student teachers simultaneously with their experiencing the subject matter education. Furthermore, MPCK can be separated into domains at this stage, which is interesting from a theoretical point of view with respect to the emergence and structure of competencies. This is important because it can help in the planning of lectures and seminars at that phase. In addition to MCK, MPCK is important for teachers, and according to the results of this article, it starts to develop together with MCK from the beginning of the training. The separation, however, shows that it is not one simple unidimensional construct and thus should be supported in addition to the mathematics lectures early in the training.

The study and the formulation of the problem arise from and are based on teacher education in Germany. Thus, the generalization to other countries might be limited. The version of the test applied here was a very first step in measuring MPCK from a content-related point of view. Now that the results have confirmed the desired possibility of separating mathematical knowledge in that framework, the test has to be expanded. In addition, testing time was limited in the study, so the test was rather short and only multiple-choice items were used. It would be desirable to include open response items for the MPCK domains. The intention is to develop a more detailed version of the instrument with the support of a broad group of experts. For further (content) validity examinations, the improved test will be conducted with students of alternative study programs like chemistry. As a benchmark examination, the improved instrument will then be applied to a) students in the second phase of teacher education and b) teachers in service. Additionally, for the quantification of criterion validity, future studies should examine how school students' competencies are affected by teachers' scores on the domains diagnostic competence and instruction competence.

References

- Ball, D. L., Lubienski, S., & Mewborn, D. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. In V. Richardson (Ed.), *Handbook of research on teaching* (Vol. 4, pp. 433–456). New York, NY: Macmillan.
- Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: what makes it special? *Journal of Teacher Education*, 59(5), 389–407. <https://doi.org/10.1177/0022487108324554>
- Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9(4), 469–520. <https://doi.org/10.1007/s11618-006-0165-2>
- Baumert, J., & Kunter, M. (2011). Das Kompetenzmodell von COACTIV. In ^[1]_[SEP]M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Eds.), *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (pp. 29 – 53). Waxmann.
- Blömeke, S., et al. (Eds.). (2011). *Kompetenzen von Lehramtsstudierenden in gering strukturierten Domänen. Erste Ergebnisse aus TEDS-LT*. Münster: Waxmann.
- Blömeke, S., Hsieh, F.-J., Kaiser, G., & Schmidt, W. H. (Eds.). (2014). *International perspectives on teacher knowledge, beliefs and opportunities to learn: TEDS-M results*. Springer.
- Blömeke, S., & Kaiser, G. (2014). Theoretical framework, study design and main results of TEDS-M. In S. Blömeke, F.-J. Hsieh, G. Kaiser, & W. H. Schmidt (Eds.), *International perspectives on teacher knowledge, beliefs and opportunities to learn: TEDS-M results*. Springer. https://doi.org/10.1007/978-94-007-6437-8_2
- 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*. Waxmann.
- Blömeke, S., Kaiser, G., & Lehmann, R. (Eds.). (2010). *TEDS-M 2008. Professionelle Kompetenz und Lerngelegenheiten angehender Mathematiklehrkräfte für die Sekundarstufe I im internationalen Vergleich*. Münster: Waxmann.
- Buchholtz, C., Doll, J., Stancel-Piatak, A., Blömeke, S., Lehmann, R., & Schwippert, K. (2011). Anlage und Durchführung der Studie TEDS-LT. In S. Blömeke et al. (Eds.), *Kompetenzen von Lehramtsstudierenden in gering strukturierten Domänen. Erste Ergebnisse aus TEDS-LT* (pp. 25 – 46). Münster: Waxmann.
- Buchholtz, N., Kaiser, G., & Stancel-Piatak, A. (2011). Professionelles Wissen von Studierenden des Lehramts Mathematik. In S. Blömeke et al. (Eds.), *Kompetenzen von Lehramtsstudierenden in gering strukturierten Domänen. Erste Ergebnisse aus TEDS-LT* (pp. 101–133). Waxmann.
- Buchholtz, N., Scheiner, T., Döhrmann, M., Suhl, U., Kaiser, G., & Blömeke, S. (2012). *TEDS-shortM: Kurzfassung der mathematischen und mathematikdidaktischen Testinstrumente aus TEDS-M, TEDS-LT und TEDS-Telekom*. Hamburg: Universität.

- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464–504. <https://doi.org/10.1080/10705510701301834>
- Cochran, K. F., DeRuiter, J. A., & King, R. A. (1993). Pedagogical content knowing: An integrative model for teacher preparation. *Journal of Teacher Education*, 44(4), 263–272. <https://doi.org/10.1177/0022487193044004004>
- 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>
- Grossman, P. L. (1990). *The making of a teacher: Teacher knowledge and teacher education*. New York: Teachers College Press.
- Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: conceptualizing and measuring teachers' topic-specific knowledge of students [Journal Article]. *Journal for Research in Mathematics Education*, 39(4), 372–400.
- Klein, F. (1908). *Elementarmathematik vom höheren Standpunkte aus. Teil 1: Arithmetik, Algebra, Analysis*. Teubner.
- Krauss, S., Brunner, M., Kunter, M., Baumert, J., Blum, W., Neubrand, M., & Jordan, A. (2008). Pedagogical content knowledge and content knowledge of secondary mathematics teachers. *Journal of Educational Psychology*, 100(3), 716 – 725. <https://doi.org/10.1037/0022-0663.100.3.716>
- Krauss, S., Neubrand, M., Blum, W., & Baumert, J. (2008). The professional knowledge of german secondary mathematics teachers: Investigations in the context of the COACTIV project. In 11th International Congress on Mathematics Education (ICME 11). Monterrey, Mexico.
- Kunter, M., & Baumert, J. (2011). Das COCATIV-Forschungsprogramm zur Untersuchung professioneller Kompetenz von Lehrkräften - Zusammenfassung und Diskussion. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Eds.), *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (pp. 345 – 366). Waxmann.
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S., & Neubrand, M. (Eds.). (2011). *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV*. Waxmann.
- Marks, R. (1990). Pedagogical content knowledge: From a mathematical case to a modified conception. *Journal of Teacher Education*, 41(3), 3–11. <https://doi.org/10.1177/002248719004100302>
- McElvany, N., Schroeder, S., Baumert, J., Schnotz, W., Horz, H., & Ullrich, M. (2012). Cognitively demanding learning materials with texts and instructional pictures: teachers' diagnostic skills, pedagogical beliefs and motivation. *European Journal of Psychology of Education*, 27(3), 403–420. <https://doi.org/10.1007/s10212-011-0078-1>

- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58(4), 525–543. <https://doi.org/10.1007/BF02294825>
- Nathan, M. J., & Petrosino, A. (2003). Expert blind spot among preservice teachers. *American Educational Research Journal*, 40(4), 905–928. <https://doi.org/10.3102/00028312040004905>
- Ohle, A., & McElvany, N. (2015). Teachers' diagnostic competences and their practical relevance. Special issue editorial. *Journal for Educational Research Online*, 7(2), 5–10.
- Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, 47(5), 667–696. <https://doi.org/10.1080/00273171.2012.715555>
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Satorra, A. (2000). Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In R. D. H. Heijmans, D. S. G. Pollock, & A. Satorra (Eds.), *Innovations in Multivariate Statistical Analysis. Advanced Studies in Theoretical and Applied Econometrics* (Vol 36). Springer, Boston, MA. https://doi.org/10.1007/978-1-4615-4603-0_17
- Schoenfeld, A. H., & Kilpatrick, J. (2008). Toward a theory of proficiency in teaching mathematics. In D. Tirosh & T. Wood (Eds.), *International handbook of mathematics teacher education* (Vol. 2, pp. 321–354). Rotherdam, Netherlands: Sense Publishers. https://doi.org/10.1163/9789087905460_016
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14. <https://doi.org/10.3102/0013189X015002004>
- Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard educational review*, 57(1), 1–23. <https://doi.org/10.17763/haer.57.1.j463w79r56455411>
- Themessl-Huber, M (2014). Evaluation of the chi²-statistic and different fit-indices under misspecified number of factors in confirmatory factor analysis. *Psychological Test and Assessment Modeling*, 56(3)
- Winsløw, C., & Grønbaek, N. (2014). The ecological Klein's double discontinuity revisited: contemporary challenges for universities preparing teachers to teach calculus. *RDM*, 34/1, 59–86.

Author Note

The project is part of the Qualitäts offensive Lehrerbildung, a joint initiative of the Federal Government and the Länder. The programme was funded by the Federal Ministry of Education and Research. Supported by Deutsche Telekom Stiftung.

Appendix

Example tasks. All tasks were applied in German.

In an exam the term

$$\int_a^b f(x)dx$$

is shown to the students (for fixed $a, b \in \mathbb{R}$). The question is what it means.
Decide which answer is correct

Mark one box per row

The term indicates...		Yes	No
A)	...always an antiderivative of f evaluated at one point.	<input type="checkbox"/>	<input type="checkbox"/>
B)	...the area between the graph of f and the x -axis, if $f(x) > 0, \forall x \in [a, b]$.	<input type="checkbox"/>	<input type="checkbox"/>
C)	...a function F with $F' = f$.	<input type="checkbox"/>	<input type="checkbox"/>

Figure A1. Example task within the schoolMCK domain

A student is asked to calculate the capacity of a car's tank (in liter). A fuel consumption of 7.6 liters per 100 km and a maximum range of 530 km are given in the task.
In seventh grade Peter gave the following, wrong answer:

After 100 km the car has used 7.6 liters. Therefore the car can go 13.16 km on one liter. Thus for the 530 km given in the task, the capacity of the tank has to be $13.16 \cdot 530 = 6974.8$ liters.

For which ones of the following tasks is it possible that Peter's mistake would also lead to a wrong answer?

Mark one box per row

		Yes	No
A)	Reduce the fraction $\frac{3x^2}{2x+8x^3}$ completely.	<input type="checkbox"/>	<input type="checkbox"/>
B)	How much is 13 % of 120€?	<input type="checkbox"/>	<input type="checkbox"/>
C)	Give all $x \in \mathbb{R}$ which solve the equation $5.6x - 12 = 0$.	<input type="checkbox"/>	<input type="checkbox"/>

Figure A2. Example task within the diagnostic competence domain. Covariances were allowed between items A and C.

A student is given the task:

Simplify the algebraic expression

$$\frac{xy+xz+yz+z^2}{y^2-z^2}$$

as much as possible.

Name possible deficits of the student.

Mark one box per row

		Yes	No
A)	The student is not able to sufficiently apply the distributive rule.	<input type="checkbox"/>	<input type="checkbox"/>
B)	The student does not know polynomial division.	<input type="checkbox"/>	<input type="checkbox"/>

Figure A3. Example task within the diagnostic competence domain. Covariances were allowed for both items.

The question arises why real numbers \mathbb{R} are always used in school even though irrational numbers hardly ever occur, which means that using rational numbers \mathbb{Q} should be enough. Which explanation would be appropriate?

Mark one box per row

		Yes	No
A)	This makes sense from a pedagogical point of view! It has been shown that certain calculations are easier for students in decimal notation compared to complicated fractions.	<input type="checkbox"/>	<input type="checkbox"/>
B)	In school, lengths, areas and volumes are frequently measured. This requires a number range which is in a biunique relation to the points on a line. This is provided by real numbers rather than by rational numbers.	<input type="checkbox"/>	<input type="checkbox"/>

Figure A4. Example task within the instruction domain.

The annual interest of a bank is 1 %, however $\frac{1}{12}\%$ interest is payed every month (thus resulting in compound interest after the first month).

A possible task for students could be:

- a) *How much interest will the saver receive after one year?*
- b) *How much would he receive with continuous interest? (i.e. following the transition from $\frac{1}{12}\%$ monthly, $\frac{1}{365}\%$ daily, ... to continuous)*

Which of the following comments seem appropriate to you?

Mark one box per row

		Yes	No
A)	This is a very complicated problem which includes probability theory.	<input type="checkbox"/>	<input type="checkbox"/>
B)	The continuous interest leads to the number $\pi = 3,1415\dots$	<input type="checkbox"/>	<input type="checkbox"/>
C)	Part (a) can be worked out using a formula, using only basic calculus. For part (b) in addition, one has to use the limit.	<input type="checkbox"/>	<input type="checkbox"/>

Figure A5. Example task within the instruction domain.