The Professional Knowledge of German Secondary Mathematics Teachers: Investigations in the Context of the COACTIV Project

Stefan Krauss¹, Michael Neubrand², Werner Blum¹ & Jürgen Baumert¹

¹University of Kassel, ²University of Oldenburg, ³Max Planck Institute for Human Development, Berlin

Project members: Mareike Kunter & Katrin Löwen, Max Planck Institute for Human Development, Berlin;
Martin Brunner, University of Luxembourg; Alexander Jordan, University of Bielefeld

0 Introduction

The relevance of teachers’ domain-specific knowledge for high-quality instruction has been emphasized repeatedly, particularly in the context of mathematics teaching (e.g., Ball, Lubinski, & Mewborn, 2001). However, despite the eminent role that is attributed to teachers’ professional knowledge for creating powerful learning environments, very few instruments are yet available to tap teachers’ knowledge directly (Baumert, Blum, & Neubrand, 2004). As a consequence, many questions on mathematics teachers’ knowledge, its content, structure, and the way it influences teaching and learning, remain empirically unresolved and, to some extent, unresearched.

One of the aims of the COACTIV project (Cognitive Activation in the Classroom) was to conceptualize the professional knowledge of secondary mathematics teachers and, on this basis, to construct reliable tests. Following Shulman’s (1986, 1987) taxonomy of teachers’ professional knowledge, we distinguish theoretically between pedagogical content knowledge (PCK), which is essentially the knowledge of “how to make the subject comprehensible to others,” and content knowledge (CK), which is the “deep understanding of the domain itself.” Shulman further identifies pedagogical knowledge (PK) as subject-independent knowledge of how to optimize learning situations in the classroom in general, which we do not address in this paper. CK, PCK, and PK are today considered the core categories of teacher knowledge (Baumert & Kunter, 2006; for additional categories, see, e.g., Shulman, 1987, or Brunner et al., 2006b).

In this paper, we report key findings concerning the tests have been constructed to assess pedagogical content knowledge and content knowledge within the COACTIV project. In so doing, we condense and summarize information presented in Krauss et al. (in press, submitted). We start by introducing the COACTIV project and its technical and conceptual framework of test construction and administration.
The COACTIV project on “Professional Competence of Teachers, Cognitively Activating Instruction, and the Development of Students’ Mathematical Literacy” was initiated with the aim of conceptualizing and assessing a broad spectrum of teacher competencies, personality variables, and work-related variables in the context of secondary mathematics instruction. The project (directors: Jürgen Baumert, Berlin; Werner Blum, Kassel; Michael Neubrand, Oldenburg) was funded by the German Research Foundation (DFG) from 2002 to 2006 and surveyed the mathematics teachers whose classes participated in the PISA 2003/2004 longitudinal assessment in Germany (see Prenzel et al., 2004, for details of PISA 2003 and its German additional option, and Prenzel et al., 2006, for details of the longitudinal German PISA component).

The close relationship between COACTIV and PISA allows representative large-scale data on teachers, their lessons, and their students to be analyzed within a common technical and conceptual framework for the first time in Germany (Figure 1). Whereas students’ achievement and personality variables were assessed in PISA (right column), their teachers were surveyed in COACTIV (left column). Parallel questionnaires on lessons (middle column) were administered to both the students (in PISA) and the teachers (in COACTIV) (“multi-

---

1 The empirical validation study of the COACTIV constructs of PCK and CK which is presently being conducted by the first author (see also section 4) is also funded by the DFG (2007-2009) and can be considered an extension of the COACTIV project.
perspectivity”). Note that Figure 1 depicts only a fraction of the constructs assessed. Besides knowledge tests, a broad battery of newly developed (or adapted) instruments tapped teachers’ biographical variables, motivational orientations, professional beliefs, and self-regulation (for an overview of the COACTIV instruments, see, e.g., Brunner et al., 2006b; Krauss et al., 2004; Kunter et al., 2007).

In combination with the data on the students obtained in the PISA study, the structure of the data allows us to use structural equation modeling to test various causal hypotheses, based on the assumption that the teacher influences lessons, which in turn influence student learning. For a general overview of the COACTIV findings, see Kunter et al. (2007) or Brunner et al. (2006b). A comprehensive reference list of COACTIV publications on specific issues (e.g., lessons in the PISA classes from the student and the teacher perspective, stress and burnout among the COACTIV teachers, teacher enthusiasm, teacher beliefs, and the mathematical tasks used in lessons) is provided in Krauss, Baumert, and Blum (submitted). In the following, we introduce the core instruments used in COACTIV, namely the tests of secondary mathematics teachers’ PCK and CK.

2 PCK and CK: Conceptualization and Test Construction

2.1 Pedagogical Content Knowledge (PCK)

Shulman (1986) characterizes pedagogical content knowledge as the knowledge needed “to make content comprehensible to others” (p. 9). Taking this as the underlying definition of PCK, we identified three subdimensions that are specifically important to mathematics teaching and used these subdimensions to guide test construction.

(1) Tasks play a central role in teaching mathematics; much of the time allocated to mathematics lessons is devoted to tasks and their solution. Appropriately selected and implemented mathematical tasks lay the foundations for students’ construction of knowledge and represent powerful learning opportunities (e.g., Jordan et al., in press). This potential can be fruitfully exploited by having students consider multiple solutions to specific problems (see, e.g., the TIMSS Video study; Neubrand, 2006). We therefore assessed teachers’ knowledge about tasks by testing their ability to produce multiple solutions of given tasks.

(2) Teachers need to work with students’ existing conceptions and prior knowledge. Because mistakes can provide valuable insights into the implicit knowledge of the problem solver (Matz, 1982), it is important for teachers to be aware of typical student misconceptions and difficulties. In our PCK test, this aspect was assessed by presenting teachers scenarios and
asking them to detect, analyze (e.g., give cognitive reasons for a given problem), or predict a
typical student error or comprehension difficulty.

(3) Students’ construction of knowledge is often only successful with appropriate
instructional support and guidance; for example, in the form of explanations or
representations. In our PCK test, teachers had to explain certain mathematical contents or
provide useful representations, analogies, illustrations, or examples to make mathematical
content accessible to students (see Kirsch, 1976).

Thus, our PCK test contained three subscales: knowledge of the multiple solution paths of
mathematical tasks (Tasks: 4 items), knowledge of student misconceptions and difficulties
(Students: 7 items), and knowledge of mathematics-specific instructional strategies
(Instruction: 11 items). Sample items from each PCK-subscale are provided in the Appendix.

2.2 Content Knowledge (CK)

Content knowledge describes a teacher’s understanding of the structures of his or her subject.
According to Shulman (1986), “the teacher need not only understand that something is so, the
teacher must further understand why it is so.” Clearly, teachers’ knowledge of the
mathematical content covered in the school curriculum should be deeper than that of their
students. Thirteen items were constructed to tap teachers’ conceptual or procedural skills in
relevant content areas (e.g., arithmetic, algebra, and geometry; see the Appendix for a sample
items). No subfacets of CK were assumed (see Krauss et al., in press).

Note that this conceptualization clearly distinguishes CK from other possible notions of
“content knowledge”: (1) the everyday mathematical knowledge that all adults should have,
(2) the school-level mathematical knowledge that school students ought to have, and (3) the
scholary mathematical knowledge that is taught at university and that does not overlap with
the content of the school curriculum (e.g., Galois theory or functional analysis). CK as
conceptualized in COACTIV lies between (2) and (3) and can best be characterized as
“elementary mathematics from a higher viewpoint” (Klein, 1933).

2.3 Test Implementation in COACTIV: Sample and Procedure

The teachers participating in COACTIV taught mathematics in the 10th grade classes within
the framework of PISA 2003/2004 in Germany. Our teacher sample thus can be considered
fairly representative of German 10th grade mathematics teachers (because the lower track
Hauptschule ends with grade 9, however, there were no Hauptschule teachers in the sample).
The COACTIV instruments were administered at two measurement points corresponding to
the dates of the German PISA assessments in April 2003 (9th grade) and April 2004 (same classes; 10th grade), with teachers being tested in the afternoon of the day their PISA students were tested. A total of 218 secondary mathematics teachers participated at the second COACTIV measurement point, when the tests of PCK and CK were implemented; 198 teachers completed both tests.

For several of the subsequent analyses, these 198 teachers were split into two groups. In Germany, the university training provided for teacher candidates differs dramatically depending on the secondary school track in which they aspire to teach. Those aspiring to teach mathematics in the academic track (Gymnasium) study the subject at a much deeper level in the first phase of their training, comparable to students majoring in mathematics.

Of the 198 teachers, 85 (55% male) taught at the academic-track Gymnasium (GY), and 113 (43% male) at other secondary school types (non-Gymnasium, NGY). One reason for subsuming the teachers of other secondary school types (mostly Realschule, but also Gesamtschule, Sekundarschule, etc.) to a single group is that the numbers in each individual group are relatively small. Furthermore, the university training provided for these other secondary school types is comparable (and substantially different from the content-focused education provided for Gymnasium teachers). The average age of the participating teachers was 47.2 years (SD = 8.4). Teachers were paid 60 euro for participation. The assessment of PCK and CK was conducted individually in a separate room at the teacher’s school on a workday afternoon. It was administered by a trained test administrator, as a power test with no time constraints. The teachers were not allowed to use a calculator. The average time required to complete the 35 items was about 2 hours (approx. 65 min for the 22 PCK items and 55 min for the 13 CK items). In terms of face validity, the teachers’ evaluation of the relevance of the items was positive (e.g., one teacher wrote: “I know I should know this”).

All 35 items were open ended. A scoring scheme was developed and 8 raters were given extensive training. Responses to each test item were coded by two raters independently. The interrater objectivity \( \rho \) was very satisfactory (on average across all items, \( \rho \) was .81). Furthermore, both tests yielded satisfactory reliabilities (Cronbach’s alpha was .78 for PCK and .83 for CK). Thus, in terms of objectivity and reliability, the test construction can be considered successful. In the following, we review the main results.
3 Results

3.1 Means and School Type Differences

The largest source of variance in teachers’ performance was whether or not they taught at a Gymnasium. As shown in Table 1, there were very large differences\(^2\) in CK \((d = 1.73)\) and large differences in PCK \((d = .80)\) with respect to school type, both indicating higher expertise in Gymnasium teachers.

Table 1: CK and PCK: Means \(M\) (Standard Deviations \(SD\)) and Empirical Maxima by Teacher Group

<table>
<thead>
<tr>
<th></th>
<th>M (SD) GY (N = 85)</th>
<th>M (SD) NGY (N = 113)</th>
<th>Effect Size (d) (GY vs. NGY)</th>
<th>Emp. Max. GY</th>
<th>Emp. Max. NGY</th>
</tr>
</thead>
<tbody>
<tr>
<td>CK (13 items)</td>
<td>8.5 (2.3)</td>
<td>4 (2.8)</td>
<td>1.73</td>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td>PCK (22 items)</td>
<td>22.6 (5.9)</td>
<td>18 (5.6)</td>
<td>0.80</td>
<td>37</td>
<td>29</td>
</tr>
<tr>
<td>Instruction (11 items)</td>
<td>9.3 (3.4)</td>
<td>7.1 (3.2)</td>
<td>0.67</td>
<td>17</td>
<td>15</td>
</tr>
<tr>
<td>Students (7 items)</td>
<td>5.8 (2.3)</td>
<td>4.3 (1.9)</td>
<td>0.71</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td>Tasks (4 items)</td>
<td>7.5 (1.8)</td>
<td>6.6 (2.0)</td>
<td>0.47</td>
<td>12</td>
<td>10</td>
</tr>
</tbody>
</table>

Note. GY = Gymnasium teachers; NGY = non-Gymnasium teachers. All differences are significant at \(p < .01\).

The large difference in CK reflects the intensive coverage of mathematical subject knowledge in GY teachers’ university training (it cannot be explained by the slightly higher cognitive ability of GY teachers; see Krauss et al., in press). However, their higher PCK scores, especially on the Student and Instruction subscales, are remarkable given that GY teachers usually receive less training in the teaching of the subject (Fachdidaktik) and in pedagogy at university. Yet this finding is in line with the results of many qualitative studies (e.g., Baumert & Kunter, 2006) that point to a close relationship between PCK and CK (see also section 3.2). Finally, as shown by Brunner et al. (2006a), when CK is statistically controlled (i.e., when only teachers with the same CK level are compared), the NGY teachers slightly outperform the GY teachers in terms of PCK.

3.2 Relationship Between PCK and CK

The relationship between the two knowledge categories can be examined directly by calculating the bivariate correlation between PCK and CK, which in the COACTIV data was .60. Note that this connection was much stronger in the GY group; indeed, modeling PCK and CK as latent constructs led to a latent correlation in the GY group that was no longer statistically distinguishable from 1 (see Krauss et al., in press). Despite this high correlation,

\(^2\) According to Cohen (1992), \(d = .20\) is a small effect, \(d = .50\) a medium effect, and \(d = .80\) a large effect.
however, the effect sizes between the two groups of teachers with respect to the two knowledge categories differed markedly ($d = 1.79$ for CK vs. “only” 0.80 for PCK).

Why was this correlation less strong in the NGY group? Closer inspection of the teacher data revealed that some NGY teachers who performed very poorly on CK (e.g., scoring only 1–2 points) nevertheless showed above-average performance on PCK. In other words, although our data support the claim that PCK profits from a solid base of CK, CK is only one possible route to PCK. The greater emphasis on pedagogy and didactics in the initial training provided for NGY teacher candidates in Germany may be another route.

3.3 Knowledge and Working Experience

Interestingly, positive correlations were not found between either of the knowledge categories and years of professional experience as a teacher (see Brunner et al., 2006a; Krauss et al., in press). These findings indicate that teachers’ knowledge no longer seems to develop a great deal once they have completed their training. This finding seems surprising; it contradicts theories that attribute teachers’ expertise development explicitly to their practical experience (Hiebert, Gallimore, & Stigler, 2002). According to “deliberate practice theory” (Ericcson, Krampe, & Tesch-Römer, 1993), however, expertise does not increase simply by doing a job. Rather, motivation and deliberate practice is required to identify and overcome one’s weaknesses, preferably with the support of ongoing expert feedback. Given that these conditions are normally not given in everyday school life (in contrast to teacher training, see below), our findings are in line with deliberate practice theory, the predictions of which have already been verified for various domains (e.g., music, sports, medicine, chess, etc.).

3.4 Knowledge and Subjective Beliefs

Analyses by Kunter et al. (2007) show that the relations of PCK and CK with teachers’ subjective beliefs on the nature of mathematics and on the learning of mathematics were in line with the expectations. For example, teachers with high PCK and CK scores tended to disagree with the view that mathematics is “just” a toolbox of facts and rules that “simply” have to be recalled and applied. Rather, these teachers tended to think of mathematics as a process permanently leading to new discoveries. At the same time, they rejected a receptive view of learning (“mathematics can best be learned by careful listening”), but tended to think that mathematics should be learned by self-determined, independent activities that foster real insight. The relationships between knowledge and beliefs fit into the desirable “profile” (Sternberg & Horvarth, 1995) of an “expert teacher” (Palmer, Stough, Burdenski, & Gonzales, 2005).
3.5 Knowledge and student learning progress

Because COACTIV was “docked onto” the PISA study, it was possible to relate teachers’ PCK to their students’ mathematics achievement gains over the year under investigation. Very briefly, when their mathematics achievement in grade 9 was kept constant, students taught by teachers with higher PCK scores performed significantly better in mathematics in grade 10 (for details, see Baumert et al., 2006). In a structural equation model, three mediators could be identified, namely the cognitive level of the mathematical tasks, classroom management, and the personal support that the teacher gave the students. Thus, PCK can explain students’ achievement gains in a non-trivial way; this finding is certainly a very important result of our studies.

Because these relations were much weaker for CK, our results demonstrate that PCK is indeed a necessary prerequisite for teachers being able to create powerful learning environments that support their students’ knowledge construction. Because student learning can be considered the ultimate aim of teaching, this finding is a strong indicator of the (predictive) validity of PCK as conceptualized and operationalized in COACTIV.

4 Construct Validation by Reference to Contrast Populations

Construct validity – i.e., the extent to which an operationalization measures the concept it purports to measure – is a crucial issue in psychometric research (see, e.g., Messick, 1988). Other criteria indicating psychometric test quality (i.e., reliability or objectivity) cannot inform on the meaning of the constructs measured; in the worst case scenario, interpretations of and conclusions drawn from test results may therefore be invalid. Evidence is thus needed to confirm that the tests of PCK and CK indeed measure “pedagogical content knowledge” and “content knowledge” (and not, for example, pedagogical knowledge or general intelligence). Since establishing validity means “collecting evidence,” it differs in this respect from reliability, which can often be calculated and expressed simply by the Cronbach’s α coefficient (e.g., Nunally & Bernstein, 1994).

In an additional construct validation study, we examined the performance of theoretically specified contrast populations on our tests of PCK and CK (for a detailed account of this study, see Krauss et al., submitted). The rationale behind this approach to construct validity was as follows: if the COACTIV tests indeed measure secondary mathematics teachers’ pedagogical content knowledge and content knowledge, it should be possible to formulate hypotheses regarding the performance of other populations on these tests. For
example, teachers of biology and chemistry (sample 2) can be expected to score rather low on both tests (especially on the CK test), whereas mathematics subject matter specialists (sample 3) can be expected to score relatively high on CK, but relatively low on PCK. At the same time, knowledge of both areas can be expected to increase continuously during teacher education. Therefore, mathematics teacher candidates (sample 4) can be expected to score higher than (even advanced) school students (sample 5) on both knowledge categories, but lower than the in-service COACTIV teachers (sample 1 in the following). In the following sections, we briefly introduce samples 2-5 (for details, see Krauss et al., submitted).

4.1 Samples and Procedure

Sample 2: Biology/chemistry teachers (Gymnasium)
Physics teaching is clearly the profession most closely related to mathematics teaching. However, it is hard to find teachers of physics who are not also teachers of mathematics. On the other hand, teachers of languages (or music, arts, religion, etc.) would probably not be able to solve the mathematics items. It therefore seemed reasonable to choose other science teachers, namely teachers of biology and chemistry, whose university training covered some aspects of mathematics. We chose teachers in Gymnasium schools who had studied and taught both biology and chemistry.

In total, 16 biology and chemistry teachers from different Berlin Gymnasium schools (all of whom were trained in and taught both biology and chemistry) were administered the COACTIV tests of PCK and CK; 12 (75%) were female and their average age was 49.1 years (SD: 6.9).

Sample 3: Students majoring in mathematics
Because professional mathematicians work in various fields (e.g., research, industry, insurance companies), the professional development of their knowledge after university is highly variable. We therefore chose to investigate students majoring in mathematics toward the end of their university career. These students are also easier to recruit and to examine in groups than professional mathematicians. Furthermore, it is possible to analyze the direct impact of their university training (without the influence of subsequent professional experience) on their PCK and CK scores.

We hypothesized that mathematics students’ CK scores would be comparable to those of Gymnasium teachers, but that their PCK scores would be considerably lower. Given the
particularly strong correlation between the two knowledge categories found for Gymnasium teachers (section 3.2), however, the mathematics students might equally be expected to score high on PCK as well.

A sample of 137 students majoring in mathematics was recruited from 7 German universities. Of the participating students, 87 (63.5%) were male and the average age was 23.9 years (SD: 1.9). On average, they had been enrolled at university for 6.4 semesters (SD: 1.9).

Sample 4: Students in advanced upper secondary mathematics courses
We investigated 18–19-year-old school students enrolled in an advanced mathematics course in grade 13 (Leistungskurs). This kind of course only exists in the Gymnasium track, where students can specialize in certain subjects in the upper secondary years. Note that these participants have not yet entered university.

The PCK and CK instruments were administered to 30 students enrolled in advanced mathematics courses in three Berlin Gymnasium schools. Of the students, 20 (67%) were male and the average age was 18.6 years (SD: 0.7).

Sample 5: Mathematics teacher candidates (Gymnasium track)
In Germany, the university training provided for students aspiring to teach at the academic track Gymnasium is comparable$^3$ to that provided for subject matter students (sample 3: students majoring in mathematics), at least in the first half of their studies. In addition, Gymnasium teacher candidates also have to study a second subject (mathematics teacher candidates often choose physics). As in sample 3 (students majoring in mathematics), we chose teacher candidates approaching the end of their university career.

A sample of 90 teacher candidates aspiring to teach mathematics at a German Gymnasium were recruited from 6 German universities. Of the teacher candidates, 37 (41%) were male and the average age was 25.2 years (SD: 2.2). On average, they had been enrolled at university for 7.7 semesters (SD: 2.4).

4.2 Results
Table 2 summarizes the performances of samples 1-5 on the tests of PCK and CK.

---

$^3$ There is some variance between the German states. In all states, however, students aspiring to teach at the academic track Gymnasium have to study the subject matter in considerably more depth than their peers aspiring to teach at the other secondary school types.
Table 2: PCK and CK: Means $M$ (and Standard Deviations $SD$) For All Samples

<table>
<thead>
<tr>
<th>Sample Nr.</th>
<th>N</th>
<th>Pedagogical Content Knowledge $M$ ($SD$)</th>
<th>Content Knowledge $M$ ($SD$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>198</td>
<td>18.6 (5.6)</td>
<td>5.9 (3.4)</td>
</tr>
<tr>
<td>$1_{GY}$</td>
<td>85</td>
<td>21.0 (5.3)</td>
<td>8.5 (2.3)</td>
</tr>
<tr>
<td>$1_{NGY}$</td>
<td>113</td>
<td>16.8 (5.1)</td>
<td>4.0 (2.8)</td>
</tr>
<tr>
<td>2</td>
<td>16</td>
<td>7.6 (2.5)</td>
<td>0.4 (0.6)</td>
</tr>
<tr>
<td>3</td>
<td>137</td>
<td>19.7 (5.1)</td>
<td>8.6 (3.0)</td>
</tr>
<tr>
<td>4</td>
<td>30</td>
<td>9.7 (5.6)</td>
<td>2.6 (2.3)</td>
</tr>
<tr>
<td>5</td>
<td>90</td>
<td>18.2 (5.0)</td>
<td>6.6 (2.8)</td>
</tr>
</tbody>
</table>

Note that the results for the COACTIV teachers are slightly different from those presented in Table 1 because two items in the PCK Instruction subscale were assessed using computer-based measures (geometrical animations were displayed). These items were not administered to the other samples for logistical reasons; the COACTIV teachers’ scores on these items were therefore excluded from the results displayed in Table 2 (this does not substantially influence any of our findings reported above). Table 2 reveals that all relationships – with the exception of the mathematics students’ mean PCK score of 19.7 – were consistent with our expectations. This exception, which is discussed in detail in Krauss et al. (submitted), can be explained at least partially by the fact that the mathematics majors who volunteered to participate in the study were probably a very selective sample who expected to be able to solve the “pedagogical content knowledge” items they had been told would be administered. The close connection between the two knowledge constructs established above (see section 3.2) – especially for the GY teachers – may further explain this striking result: content knowledge indeed seems to be a powerful source of pedagogical content knowledge. The results for the other samples are in accordance with the a priori formulated hypotheses: samples 2 and 4 scored considerably lower than the COACTIV teachers. The scores achieved by sample 5, the GY teacher candidates, were – as expected – between those attained by sample 4 and sample $1_{GY}$. Taken together, Table 2 indicates that the COACTIV knowledge tests indeed do measure something that is closely connected to the profession of mathematics teachers (and not something else, such as general intelligence).

5 Discussion

In previous studies, most conclusions about the nature of teachers’ knowledge have been drawn using indicators that are rather distal, such as university grades, number of subject
matter courses taken at university, or questionnaire data on beliefs or subjective theories (for an exception, see the research on elementary teachers by Hill, Schilling, & Ball, 2004). Consequently, numerous calls have been made in the literature for more valid and reliable assessments of teacher knowledge (e.g., Lanahan, Scotchmer, & McLaughlin, 2004). In the present paper, we have reported on the construction and implementation of tests to assess the pedagogical content knowledge and the content knowledge of secondary mathematics teachers directly. Both knowledge categories were measured reliably and the mean differences between teachers with different educational backgrounds provided evidence for the empirical validity of the tests. An additional construct validation study also supported the validity of both knowledge constructs, but yielded an unexpectedly high PCK for subject matter experts (which is analyzed in detail in Krauss et al., submitted).

Practically, our results have at least two implications. First, our instrument might find more widespread application as a psychometric assessment tool that measures teachers’ competence directly. In the light of recent developments in the areas of teacher education, selection, and accountability, this aspect is of increasing importance. Our research identifies a way of gauging teacher qualifications in terms of the assets that seem most important for their primary task of teaching. We certainly do not yet know enough about issues such as retest reliability or suitability for other samples, but addressing these questions is an important objective of our ongoing research agenda.

Second, our study provides some valuable insights into the “long arm” of university teacher training. Because no positive correlation was found between years of teaching practice and the two knowledge categories, teacher training can be assumed to be at the core of the development of the two knowledge categories. Thus, our results support current efforts to improve teacher education by placing a stronger emphasis on subject-based pedagogical content knowledge. Future research may provide deeper insights into the acquisition of PCK and CK during teacher training. For instance, longitudinal implementation of our tests at several critical stages in teacher education might provide more accurate information on the timing (e.g., in which phases of teacher education are PCK and CK acquired?) and mechanisms (e.g., which is needed to acquire the other?) of professional expertise development. Such studies may have consequences for instructional programs (at university and in the classroom) designed to foster the CK and PCK of student teachers.

Last but not least, it is our hope that the COACTIV results might not only activate the discussion on the professional knowledge of mathematics teachers, but also initiate similar endeavours for other school subjects.
5 References


Klein, F. (1933). *Elementarmathematik vom höheren Standpunkte aus* [Elementary mathematics from a higher viewpoint]. Berlin: Springer.


Appendix: Sample items and responses scoring 1 for the COACTIV tests on PCK and CK

<table>
<thead>
<tr>
<th>Knowledge Category (Subscale)</th>
<th>Sample Item</th>
<th>Sample response (scoring 1)</th>
</tr>
</thead>
</table>
| **PCK “Task”**              | How does the surface area of a square change when the side length is tripled? Show your reasoning. | **Algebraic response**  
Area of original square: \( a^2 \)  
Area of new square is then \((3a)^2 = 3^2a^2 = 9a^2\); i.e., 9 times the area of the original square.  
**Geometric response**  
Nine times the area of the original square |

| **PCK “Student”**           | The area of a parallelogram can be calculated by multiplying the length of its base by its altitude. | Note: The crucial aspect to be covered in this teacher response is that students might run into problems if the foot of the altitude is outside a given parallelogram. |

| **PCK “Instruction”**       | A student says: I don’t understand why \((-1) \cdot (-1) = 1\)  
Please outline as many different ways as possible of explaining this mathematical fact to your student. | The “permanence principle,” although it does not prove the statement, is one way to illustrate the logic behind the multiplication of two negative numbers:  
\[
\begin{align*}
3 \cdot (-1) &= -3 \\
2 \cdot (-1) &= -2 \\
1 \cdot (-1) &= -1 \\
0 \cdot (-1) &= 0 \\
(-1) \cdot (-1) &= 1 \\
(-2) \cdot (-1) &= 2
\end{align*}
\] |

| **CK**                     | Is it true that \(0.999999... = 1\)?  
Please give detailed reasons for your answer. | One possibility: Let \(0.999... = a\)  
Then \(10a = 9.99...\), hence,  
\[
10a - a = 9.99... - 0.999... \\
9a = 9 \\
Therefore a = 1; hence, the statement is true |