Computers & Education 217 (2024) 105040

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

Unraveling TPACK: Investigating the inherent structure of TPACK from a subject-specific angle using test-based instruments

Armin Fabian [a,*], Tim Fütterer [a], Iris Backfisch [b], Erika Lunowa [c], Walther Paravicini [d], Nicolas Hübner [e], Andreas Lachner [b]

[a] *University of Tübingen, Europastr. 6, 72072, Tübingen, Germany*
[b] *University of Tübingen, Wilhelmstraße 31, 72074, Tübingen, Germany*
[c] *University of Education Karlsruhe, Bismarckstraße 10, 76133, Karlsruhe, Germany*
[d] *University of Tübingen, Auf der Morgenstelle 10, 72076, Tübingen, Germany*
[e] *University of Tübingen, Münzgasse 26, 72070, Tübingen, Germany*

A R T I C L E I N F O

Keywords:
TPACK
Professional knowledge
Technology integration
Pre-service teacher education
Test-based instruments

A B S T R A C T

Against the backdrop of digitalization, it is imperative to provide pre-service teachers with adequate training opportunities to foster their professional knowledge regarding technology integration in teaching-learning scenarios. However, to date, only limited insights into the empirical nature of such knowledge – often subsumed under the term Technological Pedagogical and Content Knowledge (i.e., TPCK) – are possible given the heterogeneity of prior research investigating the empirical relationship between different knowledge components. This heterogeneity is likely due to the predominant use of self-reports in previous studies. Against this background, the present study pursued two goals. The first goal was to investigate the empirical nature of TPCK among pre-service teachers, utilizing test-based instruments to explore TPCK's nature from a subject-specific angle, that is, its relationship with Pedagogical Content Knowledge (PCK) and Technological Knowledge (TK). Given the widespread use of self-reports, the study's second goal was to examine the relationship between test-based and self-reported TPCK, exploring possible associated factors (e.g., pre-service teachers' gender, prior experience in teaching with technologies in school, or metacognitive accuracy) that may explain why both measures are linked only weakly. Findings reveal that both PCK and TK statistically predicted TPCK to a similar extent highlighting the integrated nature of TPCK. The relationship between test-based and self-reported TPCK was moderated by pre-service teachers' metacognitive accuracy, but not by their gender or prior experience. Together, these insights offer valuable guidance for refining teacher training regarding effective technology integration by indicating the need to target not only PCK and TK but also TPCK.

1. Introduction

Integrating technology in the classroom has been shown to be a lever for supporting students' learning (see Hillmayr et al., 2020 for

* Corresponding author.
E-mail addresses: armin.fabian@uni-tuebingen.de (A. Fabian), tim.fuetterer@uni-tuebingen.de (T. Fütterer), iris.backfisch@uni-tuebingen.de (I. Backfisch), erika.lunowa@ph-karlsruhe.de (E. Lunowa), w.Paravicini@uni-tuebingen.de (W. Paravicini), nicolas.huebner@uni-tuebingen.de (N. Hübner), andreas.lachner@uni-tuebingen.de (A. Lachner).

https://doi.org/10.1016/j.compedu.2024.105040
Received 14 December 2023; Received in revised form 13 March 2024; Accepted 15 March 2024
Available online 18 March 2024

a meta-analysis in STEM). Empirical evidence indicated a positive relationship between teachers' professional knowledge and the quality of instruction (Backfisch et al., 2020), which is in turn closely related to students' learning outcomes as highlighted in a recent model by Lachner et al. (2024). Hence, teachers' professional knowledge is considered a crucial pre-requisite to integrate technologies effectively in subject-matter teaching (Clark-Wilson et al., 2020; Knezek & Christensen, 2016; Mishra & Koehler, 2006). In this light, it is imperative to take into account professional knowledge regarding technology integration early on in teachers' careers and develop adequate training opportunities for pre-service teachers. However, despite its relevance for both the design of effective pre-service teacher training (Fabian et al., 2024) and the design of adequate assessment strategies (Willermark, 2018), the precise empirical nature of such professional knowledge–commonly subsumed under the Technological Pedagogical and Content model (TPACK; Mishra & Koehler, 2006)–is largely unclear. This is because prior studies investigating the empirical relationships between different knowledge components have produced mixed results. For example, to date, it is debated whether Technological Pedagogical Content Knowledge (i.e., TPCK[1]) as the central knowledge component of TPACK is a related but distinct knowledge component from Pedagogical Content Knowledge (PCK; Shulman, 1986) or merely a sub-facet of it (Große-Heilmann et al., 2022; Schubatzky et al., 2023).

The inconsistencies of prior studies on the empirical relationships of TPACK-components may have arisen as most empirical studies predominantly relied on self-report instruments to measure these components (Koehler et al., 2011; Willermark, 2018). Although ecologically reasonable and valuable in providing initial proxies for knowledge, self-reports require participants to *assess their knowledge accurately*, a process that has been shown to introduce potential biases in the level of TPCK concerning different variables, such as *gender* (Gómez-Trigueros & Yáñez de Aldecoa, 2021; Koh et al., 2010) or pre-service teachers' *experience* (e.g., Heine et al., 2023; Jang & Tsai, 2012). Therefore, it is commonly agreed upon that self-reports present a sub-optimal way to assess TPACK (Lachner et al., 2019; Scherer et al., 2017). Despite the widespread use of self-reports in the TPACK literature, little research has been conducted to explore empirically *why* the relationship between self-reported and test-based TPCK is commonly weak (Drummond & Sweeney, 2017; von Kotzebue, 2022).

Together, the review of the literature has shown two interrelated desiderates pertaining to the TPACK ecosystem. First, the inherent empirical nature of TPACK has remained largely unclear. Second, it is an open question why self-reported and test-based TPCK are commonly linked weakly. In the present study, we sought to shed light onto these two desiderates. Accordingly, we explored the inherent empirical nature of TPACK within pre-service teachers, using test-based instruments, and investigated its relationship to TK and PCK. Given the relevance of PCK for high teaching quality (Baumert et al., 2010), as well as the subject-specific nature of TPACK (Max et al., 2022; von Kotzebue, 2022; Voogt et al., 2013) and its context-sensitivity (Mishra, 2019), we employed the study in the context of mathematics education. Specifically, we examined whether the distinction between *mathematics-specific* TPCK and *mathematics-specific* PCK drawn by Mishra and Koehler (2006) is empirically warranted (see Krauss. et al., 2008, for related approaches on the relationship of PCK and CK).

Moreover, we examined the relationship between self-reported and test-based TPCK and explored possible boundary conditions on this relationship, such as pre-service teachers' gender, prior experience, and metacognitive accuracy (i.e., ability to accurately self-asses one's knowledge).

2. Theoretical background

2.1. Pre-service teachers' professional knowledge for technology integration

Integrating technologies during subject-specific instruction requires rich knowledge of how to exploit the affordances of technologies to support students' learning (Knezek & Christensen, 2016; Mishra & Koehler, 2006). Considering the importance of future teaching with technologies, much research has been put into conceptualizing such knowledge. The TPACK model introduced by Mishra and Koehler (2006) as an extension of Shulman's Pedagogical Content and Knowledge framework is one of the most prominent conceptualizations of professional knowledge (Brianza et al., 2023; Hew et al., 2019; Schmid et al., 2024). As such, it has been discussed as a "highly relevant and useful framework for informing educational research and practice" (Schmid et al., 2024, p. 14). Therefore, TPACK was used as the guiding framework in the present study. The TPACK model suggests that the central knowledge for subject-specific technology-enhanced teaching–i.e., Technological Pedagogical Content Knowledge (TPCK)–emerges from the complex interplay of basic knowledge components such as PCK and TK. PCK relates to knowledge about how to foster students' subject-specific learning best, including knowledge about their (mis-) conceptions of relevant content-related topics and how to adapt teaching strategies to students' individual needs to support their acquisition of conceptual knowledge best (Shulman, 1986). PCK has been discussed to be a central knowledge for the profession of teaching (Shulman, 1986) and predictive of high-quality instruction – especially in the context of mathematics (Backfisch et al., 2020; Baumert et al., 2010; Blömeke & Kaiser, 2014). Technological Knowledge (TK) refers to knowledge of technologies that, for instance, includes knowledge about operating with digital technologies comprising "standard sets of software tools such as word processors, spreadsheets, browsers, and e-mail" (Mishra & Koehler, 2006, p. 1027). In the ongoing digital transformation, TK has been considered crucial for implementing technologies into subject-matter teaching and learning (Fütterer et al., 2023; Kastorff et al., 2022).

[1] Note that in the present paper, we follow Schmid et al. (2020) and use the term TPACK when referring to the TPACK model as a whole (including all of its seven knowledge components, i.e., TK, PK, CK, TPK, TCK, PCK, & TPCK). Instead, we use TPCK when explicitly referring to the specific knowledge component within the TPACK model.

2.2. *The inherent structure of TPACK*

2.2.1. *Theoretical considerations*

A vivid debate on the theoretical nature of TPCK (i.e., the central TPACK-component) evolved which concerned the question of how TPCK is developed during pre-service teacher training (Angeli & Valanides, 2009), and ultimately how TPCK is related to the other TPACK-components (Schmid et al., 2020). Two opposing views crystallized around these questions: The *integrative* and *transformative* views (Angeli & Valanides, 2009). According to the integrative view, TPCK is not a distinct or unique body of knowledge. Instead, TPCK is closely related to *all* of the other TPACK-components (i.e., TK, PK, CK, PCK, TCK, and TPK). Therefore, high levels of TPCK should correspond to high levels of all other TPACK-components including PCK and TK (Schmid et al., 2020; von Kotzebue, 2022). In contrast, the transformative view posits TPCK as a distinct and unique knowledge component that goes "beyond simple integration, or accumulation, of the constituent knowledge bases" (Angeli & Valanides, 2009, p. 158). In this view, teachers' TPCK is shaped by the second-order components (e.g., PCK) but not directly by the first-order components such as TK (Schmid et al., 2020; von Kotzebue, 2022). Therefore, TPCK should be more closely related to PCK than to TK.

These unclear theoretical underpinnings may have shaped how researchers from different disciplines conceptualized TPCK regarding its relationship to other TPACK-components. For example, whereas some researchers have conceptualized TPCK as a unique knowledge facet (e.g., Angeli & Valanides, 2009), others considered TPCK to be a sub-facet of PCK (Große-Heilmann et al., 2022; Schubatzky et al., 2023). These different conceptualizations should be reflected in the empirical relationships of TPCK with PCK and TK: If TPCK was merely a sub-facet of PCK, it may be reasonable to suggest that TPCK and PCK are empirically very closely related. At the same time, as TK refers to knowledge theoretically distinct from PCK (Mishra & Koehler, 2006), assuming a lower relationship between TPCK and TK may be reasonable.

2.2.2. *Empirical studies on the relationships between TPCK, PCK, and TK*

Numerous studies have aimed to empirically clarify the inherent structure of TPACK to shed light on the ongoing debate regarding its inherent structure and inform practice in designing effective training opportunities for pre-service teachers. Considering the relevance of PCK and TK for effective technology integration, the relationships between TPCK, PCK, and TK have been studied extensively. However, empirical findings on the relationships of these TPACK-components were mixed. On the one hand, studies provided evidence that TPCK is statistically influenced by PCK but not by TK (Pamuk et al., 2015; Schmid et al., 2020). On the other hand, Koh et al. (2013) found that TPCK was influenced by TK but not PCK. Other studies indicated that neither PCK nor TK statistically influences TPCK (Dong et al., 2015).

Possibly, this heterogeneity arises as most studies relied on self-report measures in which participants need to assess their level of knowledge, a procedure which has been considered limited in assessing objective TPCK (Fütterer et al., 2023; Lachner et al., 2021; Willermark, 2018). Reasons for the limitations of self-reports have been prominently discussed in the past (Lachner et al., 2021; Schubatzky et al., 2023; von Kotzebue, 2022; Voogt et al., 2013) and encompass induced bias arising from social desirability or the challenge individuals face in accurately assessing their knowledge (Maderick et al., 2016; Urhahne & Wijnia, 2021). Especially beginners struggle to accurately assess their knowledge (Kruger & Dunning, 1999). Therefore, self-reports are valuable for providing proxies for the level of TPCK but they are not the optimal way to assess objective TPCK (see Lachner et al., 2021, for a thorough discussion on self-reports in the context of TPACK research).

In contrast, test-based instruments have been considered superior over self-reports in assessing knowledge as they provide objective, quantifiable measures of what individuals know or can do, reducing the subjectivity and bias inherent in self-reports. Moreover, test-based assessments can be systematically standardized and validated, ensuring reliability and comparability across different contexts and populations.

Given these advantages, researchers started developing and employing test-based instruments to explore the structure of TPACK from a more cognitive angle (Baier & Kunter, 2020). Most of these studies concentrated merely on subcomponents of the TPACK model, such as technology-unrelated PCK (Kilian et al., 2021; Krauss et al., 2008) or subject-unspecific TPK (Baier & Kunter, 2020; Lachner et al., 2019) and therefore did not provide insights into the relationship of TPCK with other TPACK-components. In particular, studies on the relationship between TPCK and PCK are lacking. This is surprising given the central role of PCK in the teacher profession (Baumert et al., 2010; Shulman, 1986).

A scarce example applying a test-based TPCK instrument was provided by von Kotzebue (2022) who investigated the relationship between TPCK, TPK, and TCK within a study implemented for 206 biology pre-service teachers. The author developed a new instrument to measure biology-specific TPCK which included the use of text-based vignettes depicting common teaching scenarios in biology classes. This TPCK test was employed together with test-based measures for each of the other T-components of the TPACK model which included–next to TK and TPK–Technological Content Knowledge (i.e., TCK), that is knowledge about technologies used in specific domains regardless of pedagogical applications (Mishra & Koehler, 2006). On the one hand, the results suggested that TPCK is more closely related to TPK ($r = 0.50$) than it is to TK ($r = 0.34$). On the other hand, TCK yielded a non-statistically significant relationship with TPCK ($r = 0.01$). Moreover, the results of a subsequent structural equation modeling in which TK, TPK, and TCK were specified as predictors for TPCK revealed that both TPK and TCK predicted TPCK, but TK did not. These findings were supportive of the transformative view of TPACK given the somewhat *closer* relationship of the second-order TPACK-components (i.e., TPK and TCK) compared to first-order TK. Moreover, the predictive power of TPK on TPCK was stronger ($\beta = 0.29$) than the predictive power of TCK on TPCK ($\beta = 0.69$), suggesting that TPK and TPCK are stronger related than TCK and TPCK.

In light of the critical role of PCK in teacher education (Krauss et al., 2008) and its important role in predicting student achievement (Baumert & Kunter, 2013), the question arises where PCK would fit in. To date, there has been no research on the relationship between test-based PCK and test-based TPCK using similar models, leaving the question of the predictive power of PCK on TPCK and the empirical distinguishability of both subject-specific components unresolved.

2.3. Relationship between test-based and self-reported TPCK

As illustrated above, findings on the empirical relationship between TPCK and different TPACK-components were mixed, likely due to the distortions induced by self-report measures. Although used frequently for estimating the effectiveness of technology-related interventions (Fabian et al., 2024; Willermark, 2018), or for supporting pre-service teachers in acquiring respective knowledge (Max et al., 2022), little research has been put into comparing test-based and self-report instruments assessing TPCK (see Kastorff et al., 2022, for comparing test-based and self-reported TK). A rare exception was provided by the study of von Kotzebue (2022) who compared pre-service teachers' test-based TPCK with their self-reported TPCK. She found that both measures were linked weakly ($r = 0.09$). Similar studies regarding the relationship of test-based and self-reported TPCK measures found somewhat stronger, yet still comparatively small correlations. For example, Drummond and Sweeney (2017) developed a performance-based measure including true/false items and investigated the relationship to self-reported TPCK (Kabakci Yurdakul et al., 2012). Based on a sample of 93 pre-service teachers, the correlation was small ($r = 0.24$). However, the items used to assess TPCK were rather generic in nature and did not seem to draw upon any subject-specific knowledge (e.g., "Research suggests that technology generally motivates students to participate in the teaching and learning process"; Drummond & Sweeney, 2017, p. 935). Relatedly, another study by Baier and Kunter (2020) yielded a similar weak correlation between self-reported and performance-based knowledge assessments in the context of TPK ($r = 0.15$).

In sum, previous research suggests that test-based and self-reported TPCK were only related to a limited extent, a finding which urgently needs further inspection given the widespread use of self-report measures across the TPACK landscape (Koehler et al., 2011; Willermark, 2018). To date, little attention has been brought to investigating possible explanations for the low correspondence of both measures which was exactly the second goal of the present study.

2.3.1. Variables affecting the relationship between test-based and self-reported TPCK

To examine why test-based and self-reported TPCK are empirically linked only weakly, we included three variables which might affect this relationship: gender, prior experience and metacognitive accuracy. We specifically focused on these three variables given that those have been discussed to be associated with (pre-service) teachers' self-reported TPACK-components in prior studies (gender: Koh et al., 2010; prior experience: Heine et al., 2023; metacognitive accuracy: Abbitt, 2011) and have been discussed extensively in the TPACK literature.

Gender. Early research on computer use documented that there were decisive gender differences in favor of male teachers when it comes to attitudes regarding pre-service teachers' attitudes or use of computers detached from any pedagogical context (Kay, 2006). With the advent of technologies in classrooms, researchers started to investigate whether similar results translate to TPACK-related research. However, previous research findings in the field of TPACK were inconclusive: On the one hand, several studies could not replicate the gender differences of computer usage in the field of TPACK (Jang & Tsai, 2012; Koh & Chai, 2011). On the other hand, there is growing evidence that males tend to rate their TPCK higher than females. For example, Koh et al. (2010) investigated potential gender differences among a sample of 1185 pre-service teachers and found that men rated their knowledge of teaching with technology significantly higher than females. Relatedly, Gómez-Trigueros and Yáñez de Aldecoa (2021) found similar gender effects on self-reported TPCK based on an analysis of 914 pre-service teachers from Spanish universities. Whether these findings translate to objective TPCK remains an open question.

Prior Experience. Previous studies have shown that integrated knowledge domains such as PCK (Krauss et al., 2008) or TPK (Heine et al., 2023; Lachner et al., 2019) are sensitive to pre-service teachers' level of experience (cf. knowledge encapsulation, see Boshuizen & Schmidt, 1992). Tondeur et al. (2012) also highlighted the crucial role of authentic experiences with technology integration for the development of pre-service teachers' TPCK: "It seems to be important that pre-service teachers have the possibility to see and experience the pedagogical integration of technology in the classroom during their training experiences, by observing good examples and being able to implement such practices themselves" (p. 9). In this light, pre-service teachers who had previous experience of teaching in schools (e.g., during a practical internship) should be able to outperform those in TPCK who did not. It seems worthwhile to investigate whether TPCK–as the integration of PCK and TK–also depends on (pre-service) teachers' level of experience.

Metacognitive Accuracy. Metacognitive accuracy is a term coined within the research field of metacognition (Flavell, 1979) which describes the ability of people to assess their own knowledge accurately (Maki & McGuire, 2002). Metacognitive accuracy is often conceptualized as the extent of deviation between a person's believed/self-judged skill and his/her actual level of skill or performance. It is most commonly operationalized as the squared difference between both measures: $(X_{Performance} - X_{Judgement})^2$ (Schraw, 2009). An extensive body of evidence exists that (pre-service) teachers struggle to accurately judge their own knowledge regarding technology integration (Fütterer et al., 2023; Maderick et al., 2016; Max et al., 2022). Given that answering self-report TPCK questions requires participants to assess their own knowledge, it seems reasonable to assume that the gap between test-based and self-report TPCK measures depends on participants' level of metacognitive accuracy. For example, Abbitt (2011) argued as follows: "As with any

[2] Initially, as pre-registered, we planned on comparing the level of TPCK of pre-service teachers (of mathematics) to students of mathematics without any pedagogical background. However, due to the rattler low participation rate of mathematics students, we instead decided to solely focus on pre-service teachers to keep the participant group consistent.

self-reporting measure, the ability of the instrument to accurately represent knowledge in the TPACK domains is limited by the ability of the respondents to assess their own knowledge and respond appropriately to the survey items" (p. 291). Following the author's argumentation, the better people are able to accurately assess their own knowledge (i.e., the less deviation there is between judged and actual level of skill), the more their self-reported and test-based TPCK scores should be correlated.

2.4. The present study

The two overarching goals of the present study were (1) to shed light on the empirical inherent structure of TPCK from a subject-specific angle using test-based instruments, and (2) to investigate possible boundary conditions on the relationship between test-based and self-reported TPCK.

Regarding the first goal, we specifically examined whether the theoretical distinction between PCK and TPCK drawn by Mishra and Koehler (2006) could be empirically justified. Contrasting these findings with relationships between TK and TPCK results could yield important insights into the ongoing debate on the theoretical underpinnings of TPCK as either transformative or integrative, and the thoughtful implementation of the TPACK model for the design of effective pre-service teacher training. We formulated the following two pre-registered research questions tackling our first goal.

(RQ 1a) What are the empirical relationships between TPCK, PCK, and TK of pre-service mathematics teachers assessed by subject-specific, test-based measures?

(RQ 1b) What is the predictive power of PCK and TK on TPCK and do these predictions support the transformative or integrative view of TPCK?

Given the widespread use of self-reports in the TPACK landscape, we investigated to what extent our subject specific TPCK measure was related to self-reported TPCK. As findings from prior research were rather diverse (see 2.3), we examined possible boundary conditions on this relationship. As a result, the research questions tackling the second overarching goal were as follows:

(RQ 2a) What is the relationship between test-based and self-reported TPCK within pre-service teachers?

(RQ 2b) Is this relationship affected by gender, prior experience, or the metacognitive accuracy of pre-service teachers' self-assessed level of TPCK?

3. Method

3.1. Design and sample

We employed a cross-sectional online study with secondary mathematics pre-service teachers from a total of ten different German universities. The study was conducted from March 2022 to February 2023. In Germany, pre-service teachers are obliged to study at least two content-related subjects (e.g., Mathematics and English) besides taking courses that aim at fostering specific teaching relevant knowledge, such as PCK. Another peculiarity of the pre-service teacher training curriculum in Germany is that at some point during their training, pre-service teachers are required to spend a practical internship in schools teaching mathematics, often for the first time.

Participants were selected through self-selection sampling which is a non-probability sampling technique where participants volunteer in response to an invitation, based on specific criteria (here, being pre-service mathematics teachers; see 3.3. for details on the selection process). In total, we obtained data from 150 pre-service mathematics teachers.[2] As preregistered, we eliminated 8 participants from the dataset, who finished the survey in less than 15 min (which corresponds to less than 50% of the average time participants needed to answer the knowledge tests). Further, we eliminated data from one participant who did not answer each of the test-based TPCK item presented to him/her. Therefore, our final sample included data from $N = 141$ pre-service teachers (100 females & 41 males) who were trained to teach mathematics in German upper secondary schools. On average, they were 21.5 years old ($SD = 2.17$) and either in the Bachelor's ($n = 108$) or Master's program ($n = 33$).

To keep the survey parsimonious and to avoid fatiguing effects, we implemented a planned missing data design (Graham et al., 2006). Accordingly, the pre-service teachers randomly received only some, but not all items for each of the test-based instruments (see details in the following).

3.2. Instruments

We used previously validated instruments to measure TPCK, PCK, and TK. Each construct had been demonstrated to be one-dimensional in previous studies (test-based TPCK: Lachner et al., 2021; PCK: Buchholtz et al., 2016; TK: Fütterer et al., 2023). Given that we adapted some instruments, we initially fitted a one-dimensional model for each of our test-based constructs (i.e., test-based TPCK, PCK, TK) using Confirmatory Factor Analyses (CFA), in which the constructs were represented as the only latent variable in the respective models. After establishing one-dimensionality, we transformed the raw scores of test-based TPCK, PCK, and TK by calculating estimates for the person ability as detailed in the following.

3.2.1. Test-based instruments

3.2.1.1. Mathematics specific Technological Pedagogical Content Knowledge (TPCK). To measure TPCK, we adapted the mathematics version of the test instrument developed by Lachner et al. (2021). The test-based TPCK instrument comprised eight items and was based on open-ended items that contained text vignettes that depicted typical math-specific teaching problems in the classroom (e.g., "The Pythagorean theorem is often misused in exercises in your ninth-grade class, which can be attributed to possible misconceptions about the theorem. For example, the theorem is frequently applied to triangles without a right angle, or the formula is memorized without understanding its geometric relationship. How could you use educational technologies to help your students develop a better understanding of the Pythagorean theorem and consequently apply it correctly? Please justify your answer."). From the original test used by Lachner et al. (2021), we replaced three items to ensure content validity by having the items cover the whole range of topics in the curriculum of upper secondary schools. The one-dimensional CFA indicated good fit to our data (Root Mean Square Error of Approximation [RMSEA] < 0.001; Standardized Root Mean Square Residual [SRMR] = 0.068, Robust Comparative Fit Intex [CFI] = 1.000). As the study by Lachner et al. (2021) was an effectiveness study and did not focus on the validity of the instrument, we also ensured that our TPCK instrument was reliable and valid by piloting the instrument beforehand. In this pilot study, 16 pre-service teachers answered all eight items at two distinct measurement points spaced one week apart. To account for remembering effects, we randomized the order of the test-based TPCK items at the second measurement point. To establish (test-retest) reliability, we calculated the correlation between both measurement points, which was high, $r = 0.72$. To establish validity, we investigated the convergent validity by comparing the test-based TPCK scores with the instructional quality of lesson plans (scenario adapted by Backfisch et al., 2020) provided by the pre-service teachers at the second measurement point. Again, the correlation between both measures was high, $r = 0.64$, which demonstrated the external validity of our instrument (Schmidt-Atzert & Amelang, 2012).

Of the eight test-based TPCK items, participants randomly received five of them. To account for the intended missing values, we employed a Full Information Maximum Likelihood (FIML) method as recommended (Graham, 2012). For each item, pre-service teachers could receive 0 to 3 points (see https://osf.io/y9hxt/ for the coding manual). To ensure coding quality, two raters coded 20% of all open test-based TPCK answers independently. Since interrater reliability was good ($ICC = 0.91$), the items were split in half and each rater coded the rest of the answers of their respective items independently. For the subsequent analysis, we estimated participants' test-based TPCK by calculating (manifest) factor scores using Bartlett's method (Bartlett, 1937).

3.2.1.2. Mathematics specific pedagogical Content Knowledge (PCK). To measure PCK, we used an adapted version of the well-established instrument provided in Buchholtz et al. (2016). From the available 14 original items, we had to drop one item that required drawing and thus was not transferable from paper to online format. Another item was dropped to assure that each sub-facet (content-related and pedagogical-related perspective on PCK) contained an equal number of items. Therefore, our instrument comprised 12 multiple-choice items in which participants had to apply their PCK from both a content-related and pedagogical-related perspective (e.g., "Consider in the following cases of activities in math class, which of the two forms of performance assessment – portfolio or written test – best suits the following activities: Please indicate portfolio or written test for each of the activities: a) Working on problem-solving tasks, b) Working on tasks related to calculus of variations, c) Working on tasks of algebraic manipulations"; see Buchholtz et al., 2016, for details). A one-dimensional CFA demonstrated good fit (RMSEA = 0.033; SRMR = 0.097, CFI = 1.00) to the data. Participants randomly received 8 out of 12 items and scored one point for each item if they selected the correct answer(s). To account for the planned missing values, we specified a Rasch model. For the subsequent analysis, we estimated pre-service teachers' PCK by means of Weighted Likelihood Estimates (WLEs; Warm, 1989).

3.2.1.3. Technological Knowledge (TK). To measure TK, we adapted an instrument by Fütterer et al. (2023). The original test comprised 26 single-choice items that required to apply knowledge of operating technologies that are commonly (but not exclusively) used in schools (e.g., spreadsheets or PowerPoint). To keep the test battery as small as possible, we dropped 9 items that we considered to be least relevant for high-quality instruction in the classroom (e.g., one item being dropped asked participants to choose the appropriate way to send an e-mail). Therefore, our instrument contained 17 single-choice items of which participants received 9 randomly chosen ones (e.g., "You want your students to work on a collaborative writing document. You use a common web-based Etherpad (collaborative real-time editor; e.g., edupad.ch, ZUMpad). You want to be able to distinguish the entries of different students. What specific function do Etherpads offer you? Choose the correct response: a) I can activate the function that the author colors are visible., b) I can activate the function that the students' name abbreviations are displayed before each of their entries., c) I can activate the function that only one student writes on the pad at a time., d) I can activate the function that each student writes on his or her own pad."). Again, a one-dimensional CFA demonstrated good fit to the data (RMSEA <0.001, SRMR = 0.094, CFI = 1.00). For each item that was answered correctly, participants received one point. Like for PCK, we specified a Rasch model and calculated WLEs to estimate participants' TK.

3.2.2. Other instruments

3.2.2.1. Self-reported TPCK. To measure self-reported TPCK, we used an adapted version of the questionnaire by Schmidt et al. (2009) comprising five items based on rating scales in which participants assessed their knowledge of integrating technologies for teaching mathematics (e.g., "I can teach lessons that appropriately combine mathematics, technologies, and teaching approaches"). The scale ranged from 0 = *strongly disagree* to 4 = *strongly agree* (Cronbach's $\alpha = 0.78$). For the analysis, we estimated the average score across the five items.

3.2.2.2. Metacognitive accuracy. To measure metacognitive accuracy of participants regarding their test-based TPCK, we followed common metacognitive research and asked them to make prospective judgments of learning (JOL) to estimate how well they would perform on the subsequent TPCK test (Golke et al., 2019; Jacob et al., 2020; Maki & McGuire, 2002). To operationalize metacognitive accuracy, we calculated the *Absolute Accuracy Index* as defined in Schraw (2009) which is the squared difference between the test-based TPCK scores and JOL:

$$\text{Accuracy} = (TPCK_{JOL} - TPCK_{test-based})^2.$$

Accordingly, larger values of accuracy corresponded to less accurate estimations, whereas values close to zero indicated excellent accuracy. Given that each participant answered five test-based TPCK items (see 3.2.1), the sum score of test-based TPCK ranged from 0 to 15 (three points per answer). Therefore, the scores for accuracy ranged from 0 (perfect accuracy) to $15^2 = 225$ (extremely low accuracy).

Given that the Absolute Accuracy Index does not allow to investigate over- or under-estimation (due to the square), we calculated the Bias Index which is defined as the signed difference between the test-based TPCK scores and JOL:

$$\text{Bias Index} = TPCK_{JOL} - TPCK_{test-based}\ .$$

Accordingly, negative scores on the Bias Index indicated underestimation, positive scores overestimation, and values around zero indicated good accuracy.

3.2.2.3. Prior experience with teaching in schools. To assess prior teaching experience in schools, we used a single dichotomous item in which the participants indicated whether they had completed their practical school term (1 = *prior school experience*) or not (0 = *no prior school experience*).

3.3. Procedure

Prior to the start of the study, we obtained approval from the local ethics committee. The pre-service teachers were invited to take part in the study via email or personal invitations during lectures. Upon following the provided link, the pre-service teachers gave consent to take part in the study. At the beginning of the survey, the participants provided background information and their demographic data. Next, they answered the self-reported TPCK questions. Following this, they judged how many of the upcoming test-based TPCK items they would get correct. After the test-based TPCK questionnaire, the participants answered items that tapped their PCK and TK. There was no time constraint for the participants to work on the tests and questionnaires. On average, the participants took 34.9 ($SD = 10.68$) minutes to complete the survey. They received 10 euros upon completion of the survey.

3.4. Analyses to investigate the research questions

To investigate the empirical relationships between test-based TPCK, PCK, and TK (RQ 1a), we calculated Pearson correlation coefficients for each pair of knowledge components.

To investigate the predictive power of PCK and TK on test-based TPCK (RQ 1b), we employed a multiple linear regression model, in which we specified PCK and TK as the predictors for test-based TPCK. For more precise estimates, we controlled for gender. To explore whether the predictive power of PCK and TK differed significantly, we applied a *t*-test in which we tested the hypothesis that the difference between both regression coefficients was zero (Shrout & Yip-Bannicq, 2017).

To investigate the empirical relationship between test-based and self-reported TPCK (RQ 2a), we calculated the Pearson correlation

Table 1
Descriptive Statistics of all Variables Including Two-Tailed *t*-Test Statistics to Investigate Gender Differences on these Variables.

Variables	Whole Sample ($n = 141$)		Men ($n = 41$)		Women ($n = 100$)			
	M	*SD*	*M*	*SD*	*M*	*SD*	$t(139)$	*p*
test-based TPCK[a]	0.00	0.66	0.15	0.60	−0.06	0.68	−1.73	0.087
PCK[b]	−0.03	1.02	−0.03	0.91	−0.03	1.06	0.04	0.967
TK[b]	−0.02	1.03	0.33	1.09	−0.17	0.97	−2.68	0.008
Self-reported TPCK	3.02	0.74	3.22	0.79	2.93	0.71	−2.15	0.033
Accuracy	21.74	30.85	23.25	28.12	21.13	32.02	−0.37	0.712
Judgment Bias	3.12	3.48	3.43	3.43	2.99	3.51	−0.68	0.501
	Exp.[c]	No exp.[d]	Exp.	No exp.	Exp.	No exp.		
Prior school experience	25	116	8	33	17	83		

Note. [a] Bartlett estimates, [b] WLEs.
[c] Number of participants who completed a practical internship in school.
[d] Number of participants who did not complete a practical internship in school.

coefficient.

To examine possible influencing variables on this relationship (RQ 2b), we specified three multiple linear regression models, each of which included self-reported TPCK as an independent and test-based TPCK as the dependent variable. The three models differed regarding the moderator variable, which was gender in the first, prior experience with teaching in the second, and metacognitive accuracy in the third model.

4. Results

4.1. Descriptives and preliminary analysis

All descriptive statistics of the measures used for investigating the research questions are presented in Table 1. To explore gender effects on test-based and self-reported TPCK, TK, PCK and metacognitive accuracy, we conducted two-tailed *t*-tests. Considering the last two columns of Table 1, we found statistically significant differences between male and female pre-service teachers in our sample: Men tended to self-report their TPCK higher than women, $t(139) = -2.15$, $d = 0.37$ $p = 0.033$, Further, men demonstrated significantly higher scores regarding TK than women, $t(139) = -2.68$, $d = 0.50$, $p = 0.008$.

4.2. Research question 1: the empirical relationship between TPCK, PCK, and TK

We provide scatterplots for each pair of test-based variables in Fig. 1. As suggested by the slopes of the regression lines in the scatter plots (Fig. 1), we found a statistically significant positive relationship between PCK and test-based TPCK ($r = 0.538$, $p < 0.001$) indicating that participants who performed well in PCK tended to perform also well in test-based TPCK. Similarly, we found a statistically significant, positive relationship between TK and test-based TPCK ($r = 0.472$, $p < 0.001$). Furthermore, we found a statistically significant relationship between TK and PCK ($r = 0.359$, $p < 0.001$).

To investigate the predictive power of PCK and TK on test-based TPCK, we specified a linear regression model as detailed in section 3.4.1. The summary of the statistics can be found in Table 2. Considering the first column of Table 2, we found that the effects of both TK ($\beta = 0.321$, $p < 0.001$) and PCK ($\beta = 0.423$, $p < 0.001$) on test-based TPCK were (statistically) significantly positive. However, the effect of PCK on test-based TPCK was not significantly larger than the effect of TK on test-based TPCK ($p = 0.36$).

4.3. Research question 2: the relationship between test-based TPCK and self-reported TPCK and the effect of gender, prior experience and metacognitive accuracy

The relationship of self-reported TPCK and test-based TPCK was low but statistically significant, $r = 0.243$, 95% CI [0.081, 0.392], $p = 0.004$. To investigate possible influencing variables on this relationship, we specified three linear models with either gender, or prior experience, or metacognitive accuracy as moderators. The complete statistics are displayed in Table 3.

Across all three models, we found a statistically significant positive effect of self-reported TPCK on test-based TPCK with varying effect sizes (see Table 3) indicating that pre-service teachers who reported having higher levels of TPCK also performed better in the test-based TPCK questions.

Regarding main effects of moderators, no significant effect of gender on test-based TPCK was observed (Model 1 in Table 3). In contrast, the main effect of prior experience on test-based TPCK was significantly positive, $\beta = .590$, $p = .006$ indicating that pre-service teachers with prior teaching experience in school outperformend those without respective experience. Additionally, the

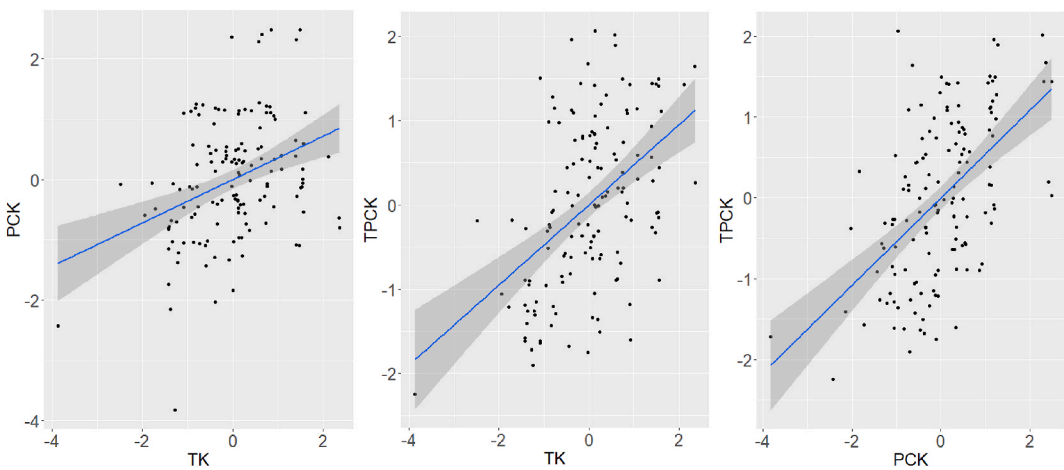


Fig. 1. Scatter Plots for Each Pair of Performance-Based Measures. *Note.* In each of the scatter plots, a dot represents the unstandardized score of one participant in the respective measures.

Table 2
Regression coefficients of PCK and TK on test-based TPCK.

Variable	β	*SE*	*t*	*p*
(Intercept)	0.000	0.007	0.00	1
TK	0.321	0.007	4.460	<0.001
PCK	0.423	0.007	5.885	<0.001
R^2_{adj}	0.37			

Note. $N = 141$. *SE* denotes the standard error. β denotes the standardized regression coefficients.

Table 3
Three Different Linear Models Displaying Regression Coefficients of Self-Reported TPCK on test-based TPCK with Gender (Model 1), Prior Experience (Model 2), or Metacognitive Accuracy (Model 3) as Moderators.

Variable	Model 1				Model 2				Model 3			
	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*
(Constant)	-0.069	0.098	−0.703	0.483	−0.116	0.088	−1.320	0.189	0.016	0.065	0.250	0.803
SR-TPCK[a]	0.204	0.102	1.991	0.048	0.175	0.087	2.019	0.045	0.321	0.066	4.880	<0.001
Moderator												
Gender	0.221	0.186	1.188	0.237								
Interaction	0.061	0.178	0.342	0.733								
Experience[b]					0.590	0.212	2.780	0.006				
Interaction					0.307	0.227	1.353	0.178				
Accuracy									−0.564	0.066	−8.600	<0.001
Interaction									−0.157	0.056	−2.815	0.006
R^2_{adj}	0.050				0.111				0.406			

Note. [a] SR-TPCK = Self-reported TPCK, [b] Experience: 0 = Pre-service teacher did not complete practical internship, 1 = Pre-service teacher completed practical internship.

main effect of accuracy on test-based TPCK was significantly negative, $\beta = -0.564$, $p < 0.001$, indicating that the more accurate pre-service teachers could assess their level of TPCK (i.e., values of accuracy tended towards zero), the better they performed in the TPCK test.

Regarding interaction effects, we found no evidence that gender or prior experience moderated the relationship between self-reported and test-based TPCK (Table 3). However, we found a statistically significant interaction effect of accuracy, $\beta = -0.157$, $p = 0.006$, which suggests that the relationship between self-reported and test-based TPCK was higher for pre-service teachers who could more accurately assess their TPCK. To explore this interaction further, we employed a subsequent Johnson-Neyman procedure (Johnson & Fay, 1950). As suggested by the Johnson-Neyman plot (Fig. 2), for pre-service teachers whose accuracy was below a

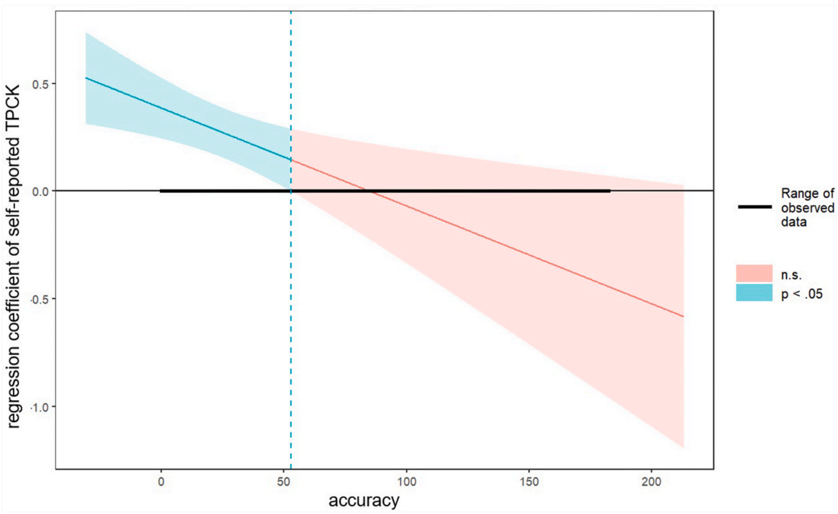


Fig. 2. Johnson-Neyman Plot for the Moderating Effect of Accuracy on the Relationship Between Test-Based and Self-Reported TPCK. *Note.* The range of observed values of accuracy was [0, 182.25]. Values are represented in unstandardized beta coefficients. When participants' accuracy was below the cut-off value of 52.84 (blue area), the effect of self-reported TPCK on test-based TPCK was statistically significant.

threshold, the relationship between test-based TPCK and self-reported TPCK was statistically significant. Put differently, the more accurate participants could self-assess their knowledge regarding technology integration, the stronger both measures (self-reported TPCK and test-based TPCK) were linked.

5. Discussion

In the present study, we shed light on two desiderate that have prevailed in the TPACK ecosystem for years. First, we investigated the inherent structure of TPACK by employing test-based and subject-specific instruments, and thereby extended prior research that have mainly focused on the relationship of subject-unspecific TPACK-components such as TPK. In particular, we were interested in scrutinizing whether the theoretical distinction between PCK and TPCK is empirically warranted. By also including TK as a predictor for TPCK, we were further able to compare the predictive power of TK on TPCK with the predictive power of PCK on TPCK to conclude whether the data are suggestive of the transformative or integrative view of TPACK. Second, we explored potential factors contributing to the predominant findings of prior research indicating that test-based and self-report TPCK instruments have only been linked to a limited extent. To do so, we investigated potential influencing variables that have been shown to be confounded with the level of self-reported TPCK (i.e., gender, prior experience, and metacognitive accuracy).

5.1. Main findings

5.1.1. Research question 1: the inherent structure of TPACK

Concerning the inherent structure of TPACK, the results suggest that TPCK and PCK are moderately related ($r = 0.538$) and therefore similar, yet different knowledge components (Angeli & Valanides, 2009). In terms of the empirical distinguishability, we can assert that TPCK represents a knowledge facet on its own rather than a sub-facet of PCK, as both constructs only share 30% of common variance. Considering the predictive power of PCK and TK on TPCK, our findings suggest that both PCK and TK moderately predicted the level of pre-service teachers' TPCK. This finding indicates that TK, too, is essential for TPCK which contrasts the findings by von Kotzebue (2022) who did not find a statistically significant effect of TK on TPCK in the context of biology education. This finding highlights the situated character of TPACK in a (domain) specific context and calls for further investigations across different domains.

The similar *closeness* of TK and PCK with TPCK highlights the importance of both constructs for domain-specific teaching with technologies. Moreover, comparing the predictive power of PCK on TPCK ($\beta = 0.423$) in our sample with predictors for TPCK from the study of von Kotzebue (2022), we further note that PCK seems to be more predictive for TPCK than TCK but less predictive than TPK. Again, one needs to be cautious as TPCK is highly context sensitive (Mishra, 2019; Schmid et al., 2020) and comparisons across different samples are hence limited. Contrasting our findings with prior studies based on self-reports, we note that we could not confirm prior findings that "TPCK is primarily influenced by the hybrid components TPK and PCK" (Schmid et al., 2020) as we found a statistically significant contribution of TK to TPCK, too. Also, our findings are in line with some studies that showed that PCK was predictive of TPCK (e.g., Celik et al., 2014; Schmid et al., 2020) while contrasting findings from other studies that instead found TCK and TPK to be predictive of TPCK (e.g., Dong et al., 2015; Koh et al., 2013).

5.1.2. Research question 2: the relationship between test-based and self-reported TPCK

Our results suggested that test-based TPCK was only linked weakly to self-reported TPCK which is in line with an ever-growing body of evidence from previous research (Drummond & Sweeney, 2017; Max et al., 2022; von Kotzebue, 2022). In contrast to prior studies, however, we also investigated boundary conditions on this relationship by including possible influencing variables, which allowed us to recognize the following patterns in our data.

5.1.2.1. The role of gender. Male pre-service teachers self-reported their TPCK abilities significantly higher than females (Table 1) which is in line with prior results from studies based on self-reports (see 2.3.1). At the same time, there was no main effect of gender on test-based TPCK. Taking both findings together, we may conclude that self-reported TPCK does induce gender-effects that stem from the way TPCK is assessed. At the same time, our moderator analyses did not provide evidence for a moderating gender effect on the relationship between test-based and self-reported TPCK.

5.1.2.2. The role of prior experience. We found a main effect of experience on test-based TPCK indicating that pre-service teachers who had had the opportunity to teach students in authentic school settings were able to integrate technology more effectively to support students' learning. This finding highlights the crucial factor of providing pre-service teachers with opportunities to gain experience and reflect on their actions as suggested by the SQD-model (Tondeur et al., 2012) and prior studies (e.g., Lachner et al., 2021). Similar to gender, we did not find a moderating effect of experience on the relationship between self-reported and test-based TPCK.

5.1.2.3. The role of metacognitive accuracy. We did not only find a significant main effect on TPCK but also a significantly moderating effect of metacognitive accuracy on TPCK suggesting that the relationship between self-reported and test-based TPCK was higher for those pre-service teachers who assessed their level of TPCK more accurately.

5.2. Theoretical implications

Our findings reveal two principal theoretical insights. First, while PCK and TPCK are moderately related, they still seem to be distinct knowledge components mirroring findings about the empirical distinguishability of PCK and CK (Krauss et al., 2008). Therefore, the theoretical distinction between PCK and TPCK seems to be warranted despite existing conceptualizations that consider TPCK a sub-facet of PCK (Schubatzky et al., 2023). At the same time, the results suggest that PCK and TK are both equally necessary for technology-enhanced instruction in subject-specific teaching. Also, PCK and TK could not explain the whole variance of TPCK indicating that TPCK goes beyond the mere integration of PCK and TK. This may be suggestive of TPCK being of transformative nature which is in line with recent findings (von Kotzebue, 2022). Second, our findings revealed that metacognitive skills play a central role in the reliability of self-report instruments. Put differently, self-report measures seem to be a suitable way to assess pre-service teachers' knowledge of technology integration if and only if their level of metacognitive accuracy is high. Against this background, developing self-report instruments including items with more context may be a promising pathway to help teachers judge their knowledge more accurately (see Sailer et al., 2021, for a first attempt in this regard).

5.3. Practical implications

What are the practical implications of our findings? First, our findings highlight the distinct role of knowledge regarding technology integration in a subject-specific setting. Seemingly, providing high quality teaching with technology is not merely dependent on Pedagogical Content Knowledge. Instead, (pre-service) teachers need to combine their knowledge of *good* mathematics teaching and learning with their knowledge of technologies. Therefore, pre-service teacher trainings should provide plenty opportunities for pre-service teachers to deliberately practice integrating PCK with TK to reach TPCK.

Second, the main effect of prior experience on TPCK indicates the need for early practical experiences in pre-service teacher education. The earlier pre-service teachers are allowed to experience authentic teaching and reflect on their acting, the better they seem to be prepared for their future life as teachers.

Third, the crucial role of metacognitive accuracy calls forth the need for not only focusing on cognitive but also on metacognitive competencies throughout pre-service teacher curricula. For example, constant formative diagnoses within lectures and seminars, or the application of diaries in which pre-service teachers reflect on their learning may pose adequate strategies to improve pre-service teachers' metacognitive accuracy (Cohen-Sayag & Fischl, 2020; Saks & Leijen, 2019).

5.4. Limitations and conclusion

One needs to consider several limitations when interpreting our findings. First, our study was of correlational nature. Therefore, causal interpretations should be avoided. For example, it remains unclear whether high levels of PCK or TK automatically translate to high levels of TPCK. Here, more research is needed with strong experimental research designs to establish causality (see Evens et al., 2018, for an experimental approach in the context of English-specific PCK). Second, we implemented the study in an online survey which carries the risk of little control and inattentiveness of participants. At the same time, the use of an online format, not requiring pre-service teachers to provide identifiable data, likely mitigated social desirability biases. Third, we employed a self-selected sampling technique, which may have introduced a volunteer bias, potentially affecting the generalizability of the findings. Fourth, our TK focused narrowly on operational aspects, omitting other crucial facets of TK, such as ethical consequences (Gómez-Trigueros, 2023) or future technologies like artificial intelligence (Celik, 2023). Fifth, we included only some, but not all, of the seven TPACK-components in our study. Possibly, some of the unexplained variance in TPCK could be accounted for by the varying levels of TCK or PK. Therefore, we encourage future researchers to conduct studies considering *all seven TPACK-components*. To implement such a study, however, researchers need to invest more resources into the development of parsimonious subject-specific test-based instruments (e.g., multiple choice tests, see Große-Heilmann et al., 2023, for first attempts in the field of physics education) to make sure not to overburden participants and to keep studies economically feasible. Lastly, our study only allows to draw conclusions regarding the inherent structure of TPACK from mathematics pre-service teachers trained at universities in Germany. Given the context sensitivity of TPACK (Mishra, 2019), it is questionable to what extent our results translate to different *subjects* or *levels of expertise*, such as in-service teachers. Therefore, future research is necessary to investigate whether our results pertain across different contexts.

Despite these limitations, our findings contribute to the nuanced understanding of the inherent structure of TPACK in a subject-specific context, and provide valuable insights into the peculiarities of different assessments.

Funding

The research reported in this article was supported by the Federal Ministry of Education and Research in Germany (BMBF) under contract number 01JA2009 and by the LEADGraduate School & Research Network, funded by the Ministry of Science, Research and the Arts of the state of Baden-Württemberg within the framework of the sustainability funding for the projects of the Excellence Initiative II.

CRediT authorship contribution statement

Armin Fabian: Writing – review & editing, Writing – original draft, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Tim Fütterer:** Writing – review & editing, Methodology, Formal analysis, Data curation. **Iris Backfisch:** Writing – review & editing, Conceptualization. **Erika Lunowa:** Writing – review & editing, Data curation. **Walther Paravicini:** Supervision, Conceptualization. **Nicolas Hübner:** Writing – review & editing, Methodology. **Andreas Lachner:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

none.

Data availability

The prepared data, analysis script and coding manual used in this study are available at the Opern Science Framework repository and can be accessed via https://osf.io/n37bp/.

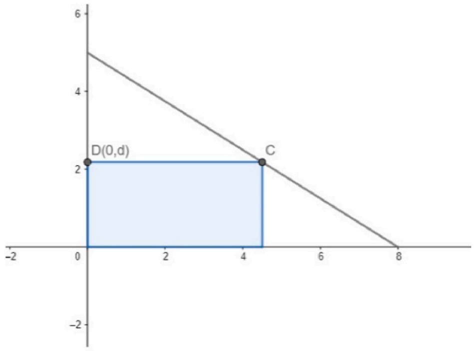
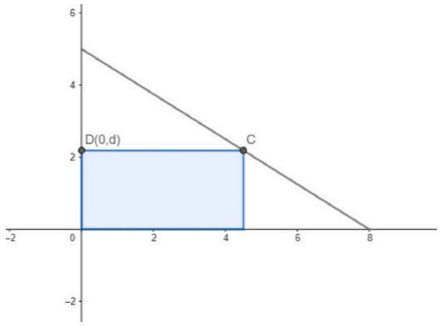
Appendix

The Test-Based TPCK Items

Items	Original Item (in German)	Translated Item
Item 1	Eine andere Lehrkraft aus Ihrem Kollegium möchte das bestimmte Integral als Grenzwert einer Summe von Rechtecken einführen ("Riemannsummen"). Sie bittet Sie dabei um Rat für eine mögliche Umsetzung in ihrem Unterricht mit digitalen Medien. *Was würden Sie der Lehrkraft empfehlen, um für ihre geplante Einführung digitale Medien gewinnbringend einzusetzen? Begründen Sie dabei gegenüber der anderen Lehrkraft Ihren Vorschlag.*	Another teacher from your department wants to introduce the concept of the definite integral as the limit of a sum of rectangles ("Riemann sums"). He/she is seeking your advice on how this could be implemented in their teaching using educational technologies. *What would you recommend to the teacher for effectively utilizing educational technologies in their planned introduction? Justify your suggestion to the other teacher.*
Item 2	Das korrekte Multiplizieren von zwei Brüchen $\left(\frac{a}{b}\cdot\frac{c}{d}=\frac{ac}{bd}\right)$ bereitet den meisten Ihrer Schülerinnen und Schüler in der sechsten Klasse nur wenig Schwierigkeiten. Allerdings fällt es vielen in Ihrer Klasse schwer zu begründen, dass das Produkt zweier rationaler Zahlen betragsmäßig kleiner sein kann als die einzelnen Faktoren, denn sie haben die Fehlvorstellung, dass durch Multiplikation zweier Zahlen stets eine betragsmäßig größere Zahl entstehen muss. *Wie könnten Sie digitale Medien einsetzen, um dieser Fehlvorstellung Ihrer Schülerinnen und Schüle entgegenzuwirken? Begründen Sie bitte Ihre Antwort.*	The correct multiplication of two fractions $\left(\frac{a}{b}\cdot\frac{c}{d}=\frac{ac}{bd}\right)$ presents little difficulty for most of your sixth-grade students. However, many in your class struggle to justify that the product of two rational numbers can be smaller in absolute value than the individual factors, as they have the misconception that multiplying two numbers must always result in a number that is larger in absolute value. *How could you use educational technologies to counteract this misconception among your students? Please justify your answer.*
Item 3	Der Satz des Pythagoras wird in Anwendungsaufgaben in Ihrer neunten Klasse häufig falsch verwendet, was auf möglichen Fehlvorstellungen des Satzes zurückzuführen ist. So wird der Satz häufig bei Dreiecken ohne rechten Winkel angewandt oder die Formel $\left(\mathrm{a}^2+\mathrm{b}^2=\mathrm{c}^2\right)$ auswendig gelernt, ohne deren geometrischen Bezug zu verstehen. *Wie könnten Sie digitale Medien einsetzen, damit Ihre Schülerinnen und Schüler ein besseres Verständnis des Satzes des Pythagoras entwickeln und ihn als Folge korrekt anwenden? Begründen Sie bitte Ihre Antwort.*	The Pythagorean theorem is often misused in exercises in your ninth-grade class, which can be attributed to possible misconceptions about the theorem. For example, the theorem is frequently applied to triangles without a right angle, or the formula is memorized without understanding its geometric relationship. *How could you use educational technologies to help your students develop a better understanding of the Pythagorean theorem and consequently apply it correctly? Please justify your answer*
Item 4	Variiert man bei einer gegebenen Funktion der Form $\mathrm{f(x)}=\mathrm{ax}^2+\mathrm{bx}+\mathrm{c}$ den Parameter *c*, so verschiebt sich der dazugehörige Graph nach oben oder unten. Auf der Grundlage dieses Wissens nehmen mehrere Ihrer Schülerinnen und Schüler in der achten Klasse an, dass sich der Graph auch durch das Variieren von a ausschließlich nach oben oder unten bewegt. *Wie könnten Sie digitale Medien verwenden, um das konzeptuelle Fehlverständnis Ihrer Schülerinnen und Schüler zu korrigieren? Begründen Sie bitte Ihre Antwort.*	When varying the parameter c in a given function of the form $f(x)=ax^2+bx+c$ the corresponding graph shifts up or down. Based on this knowledge, several of your eighth-grade students assume that varying a also moves the graph exclusively up or down. *How could you use educational technologies to correct this conceptual misunderstanding among your students? Please justify your answer.*
Item 5	Sie möchten Ihre Schülerinnen und Schüler an Problemlösestrategien bei Extremwertaufgaben unter Nebenbedingungen heranführen. Dafür legen Sie den Schülerinnen und Schüler die folgende Aufgabe vor, die sie selbstständig bearbeiten sollen: „Berechnen Sie den Parameter *d* des Punktes *D* so, dass der (blau-schraffierte) Flächeninhalt des eingeschlossenen Rechtecks (siehe Bild) maximal wird.“	You want to introduce your students to problem-solving strategies for optimization problems with constraints. For this purpose, you present the following task to your students, which they should work on independently: "Calculate the parameter *d* of point *D* such that the (blue-hatched) area of the enclosed rectangle (see image) becomes maximal."

(*continued on next page*)

(*continued*)

Items	Original Item (in German)	Translated Item
	 *Wie könnten Sie digitale Medien einsetzen, um Ihre Schülerinnen und Schüler bei der Bearbeitung der Aufgabe zu unterstützen? Begründen Sie bitte Ihre Antwort.*	 *How could you use educational technologies to support your students in solving this task? Please justify your answer.*
Item 6	Sie möchten in einem lernendenzentrierten Unterrichtsszenario den Satz des Thales in Ihrer siebten Klasse einführen. Ihre Schülerinnen und Schüler sollen dabei eigenständig eine Vermutung für die Aussage des Satzes aufstellen und Begründungen finden, die die Richtigkeit des Satzes plausibel machen. *Welche Einsatzmöglichkeiten digitaler Medien könnten Sie sich vorstellen, um ihr Vorhaben in die Tat umzusetzen? Begründen Sie bitte Ihre Antwort.*	You want to introduce Thales' theorem in a learner-centered teaching scenario in your seventh-grade class. Your students should independently formulate a hypothesis for the statement of the theorem and find justifications that make the correctness of the theorem plausible. *What potential uses of educational technologies could you envision to implement your plan? Please justify your answer.*
Item 7	Sie haben in Ihrer siebten Klasse das Lösen von Linearen Gleichungen mit einer Variablen mit Äquivalenzumformungen eingeführt. Sie möchten nun, dass Ihre Schülerinnen und Schüler diese Rechenfertigkeit daheim bis zur nächsten Stunde einüben. *Welche Einsatzmöglichkeiten digitaler Medien könnten Sie sich vorstellen, um das Einüben der Rechenfertigkeit sinnvoll zu unterstützen? Bitte begründen Sie Ihre Antwort.*	You have introduced solving linear equations with one variable using equivalence transformations in your seventh-grade class. Now, you want your students to practice this computational skill at home until the next lesson. *What potential uses of educational technologies could you envision to meaningfully support the practice of this computational skill? Please justify your answer.*
Item 8	Sie wollen Ihren Schülerinnen und Schülern in Ihrem Leistungskurs den Erwartungswert einer Zufallsvariablen als den Durchschnitt des beobachteten Werts eines sich unbegrenzt wiederholenden gleichbleibenden Experiments verdeutlichen ("Gesetz der groβen Zahlen"). *Wie könnten Sie dabei digitale Medien lernwirksam einsetzen, um den beschriebenen Zusammenhang zu verdeutlichen? Begründen Sie bitte Ihre Antwort.*	You want to demonstrate to your students in your advanced course the expected value of a random variable as the average of the observed value of an unlimited, consistently repeating experiment ("Law of Large Numbers"). *How could you effectively use educational technologies to illustrate the described relationship? Please justify your answer.*

References

Abbitt, J. T. (2011). Measuring technological pedagogical content knowledge in preservice teacher education. *Journal of Research on Technology in Education, 43*(4), 281–300. https://doi.org/10.1080/15391523.2011.10782573

Angeli, C., & Valanides, N. (2009). Epistemological and methodological issues for the conceptualization, development, and assessment of ICT–TPCK: Advances in technological pedagogical content knowledge (TPCK). *Computers & Education, 52*(1), 154–168. https://doi.org/10.1016/j.compedu.2008.07.006

Backfisch, I., Lachner, A., Hische, C., Loose, F., & Scheiter, K. (2020). Professional knowledge or motivation? Investigating the role of teachers' expertise on the quality of technology-enhanced lesson plans. *Learning and Instruction, 66*, Article 101300. https://doi.org/10.1016/j.learninstruc.2019.101300

Baier, F., & Kunter, M. (2020). Construction and validation of a test to assess (pre-service) teachers' technological pedagogical knowledge (TPK). *Studies In Educational Evaluation, 67*, Article 100936. https://doi.org/10.1016/j.stueduc.2020.100936

Bartlett, M. S. (1937). The statistical conception of mental factors. *British Journal of Psychology, 28*(1), 97–104. https://doi.org/10.1111/j.2044-8295.1937.tb00863.x. *General Section*.

Baumert, J., & Kunter, M. (2013). The COACTIV model of teachers' professional competence. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Eds.), *Cognitive activation in the mathematics classroom and professional competence of teachers: Results from the COACTIV project*. Springer US. https://doi.org/10.1007/978-1-4614-5149-5.

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss Dubberke, 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, 47.

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* (pp. 19–47). Netherlands: Springer. https://doi.org/10.1007/978-94-007-6437-8_2.

Boshuizen, H. P. A., & Schmidt, H. G. (1992). On the role of biomedical knowledge in clinical reasoning by experts, intermediates and novices. *Cognitive Science, 16*(2), 153–184. https://doi.org/10.1207/s15516709cog1602_1

Buchholtz, N., Scheiner, T., Döhrmann, M., Suhl, U., Kaiser, G., & Blömeke, S. (2016). TEDS-shortM: Teacher education and development study – short test on mathematics content knowledge (MCK) and mathematics pedagogical content knowledge (MPCK). *Kurzfassung der mathematischen und mathematikdidaktischen Testinstrumente aus TEDS-M, TEDS-LT und TEDS-Telekom*.

Celik, I. (2023). Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. *Computers in Human Behavior, 138*, Article 107468. https://doi.org/10.1016/j.chb.2022.107468

Celik, I., Sahin, I., & Akturk, A. O. (2014). Analysis of the relations among the components of technological pedagogical and content knowledge (tpack): A structural equation model. *Journal of Educational Computing Research, 51*(1), 1–22. https://doi.org/10.2190/EC.51.1.a

Clark-Wilson, A., Robutti, O., & Thomas, M. (2020). Teaching with digital technology. *ZDM - International Journal on Mathematics Education, 52*, 1223–1242. https://doi.org/10.1007/s11858-020-01196-0

Cohen-Sayag, E., & Fischl, D. (2020). Reflective writing in pre-service teachers' teaching: What does it promote? *Australian Journal of Teacher Education, 37*(10), 20–36. https://doi.org/10.3316/aeipt.198796

Dong, Y., Chai, C. S., Sang, G.-Y., Koh, J. H. L., & Tsai, C.-C. (2015). Exploring the profiles and interplays of pre-service and in-service teachers' technological pedagogical content knowledge (TPACK) in China. *Journal of Educational Technology & Society, 18*(1), 158–169.

Drummond, A., & Sweeney, T. (2017). Can an objective measure of technological pedagogical content knowledge (TPACK) supplement existing TPACK measures?: TPACK-deep and objective TPACK. *British Journal of Educational Technology, 48*. https://doi.org/10.1111/bjet.12473

Evens, M., Elen, J., Larmuseau, C., & Depaepe, F. (2018). Promoting the development of teacher professional knowledge: Integrating content and pedagogy in teacher education. *Teaching and Teacher Education, 75*, 244–258. https://doi.org/10.1016/j.tate.2018.07.001

Fabian, A., Backfisch, I., Kirchner, K., & Lachner, A. (2024). *A systematic review and meta-analysis on TPACK-based interventions from a perspective of knowledge integration* [Manuscript submitted for publication].

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist, 34*(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906

Fütterer, T., Steinhauser, R., Zitzmann, S., Scheiter, K., Lachner, A., & Stürmer, K. (2023). Development and validation of a test to assess teachers' knowledge of how to operate technology. *Computers and Education Open, 5*, Article 100152. https://doi.org/10.1016/j.caeo.2023.100152

Golke, S., Hagen, R., & Wittwer, J. (2019). Lost in narrative? The effect of informative narratives on text comprehension and metacomprehension accuracy. *Learning and Instruction, 60*, 1–19. https://doi.org/10.1016/j.learninstruc.2018.11.003

Gómez-Trigueros, I. M. (2023). Digital skills and ethical knowledge of teachers with TPACK in higher education. *Contemporary Educational Technology, 15*(2), Article ep406. https://doi.org/10.30935/cedtech/12874

Gómez-Trigueros, I. M., & Yáñez de Aldecoa, C. (2021). The digital gender gap in teacher education: The TPACK framework for the 21st century. *European Journal of Investigation in Health, Psychology and Education, 11*(4). https://doi.org/10.3390/ejihpe11040097. Article 4.

Graham, J. W. (2012). *Missing data: Analysis and design*. Springer Science & Business Media.

Graham, J. W., Taylor, B. J., Olchowski, A. E., & Cumsille, P. E. (2006). Planned missing data designs in psychological research. *Psychological Methods, 11*(4), 323–343. https://doi.org/10.1037/1082-989X.11.4.323

Große-Heilmann, R., Burde, J.-P., Riese, J., Schubatzky, T., & Weiler, D. (2023). *Erwerb und Messung fachdidaktischen Wissens zum Einsatz digitaler Medien*.

Große-Heilmann, R., Riese, J., Burde, J.-P., Schubatzky, T., & Weiler, D. (2022). Fostering pre-service physics teachers' pedagogical content knowledge regarding digital media. *Education Sciences, 12*(7). https://doi.org/10.3390/educsci12070440. Article 7.

Heine, S., Krepf, M., Jäger-Biela, D. J., Gerhard, K., Stollenwerk, R., & König, J. (2023). Preservice teachers' professional knowledge for ICT integration in the classroom: Analysing its structure and its link to teacher education. *Education and Information Technologies*. https://doi.org/10.1007/s10639-023-12212-7

Hillmayr, D., Ziernwald, L., Reinhold, F., Hofer, S. I., & Reiss, K. M. (2020). The potential of digital tools to enhance mathematics and science learning in secondary schools: A context-specific meta-analysis. *Computers & Education*, 1–25. https://doi.org/10.1016/j.compedu.2020.103897

Jacob, L., Lachner, A., & Scheiter, K. (2020). Learning by explaining orally or in written form? Text complexity matters. *Learning and Instruction, 68*, Article 101344. https://doi.org/10.1016/j.learninstruc.2020.101344

Jang, S.-J., & Tsai, M.-F. (2012). Exploring the TPACK of Taiwanese elementary mathematics and science teachers with respect to use of interactive whiteboards. *Computers & Education, 59*(2), 327–338. https://doi.org/10.1016/j.compedu.2012.02.003

Johnson, P. O., & Fay, L. C. (1950). The Johnson-Neyman technique, its theory and application. *Psychometrika, 15*(4), 349–367. https://doi.org/10.1007/BF02288864

Kabakci Yurdakul, I., Odabasi, H. F., Kilicer, K., Coklar, A. N., Birinci, G., & Kurt, A. A. (2012). The development, validity and reliability of TPACK-deep: A technological pedagogical content knowledge scale. *Computers & Education, 58*(3), 964–977. https://doi.org/10.1016/j.compedu.2011.10.012

Kastorff, T., Sailer, M., Vejvoda, J., Schultz-Pernice, F., Hartmann, V., Hertl, A., Berger, S., & Stegmann, K. (2022). Context-specificity to reduce bias in self-assessments: Comparing teachers' scenario-based self-assessment and objective assessment of technological knowledge. *Journal of Research on Technology in Education, 0*(0), 1–14. https://doi.org/10.1080/15391523.2022.2062498

Kay, R. (2006). Addressing gender differences in computer ability, attitudes and use: The laptop effect. *Journal of Educational Computing Research, 34*(2), 187–211. https://doi.org/10.2190/9BLQ-883Y-XQMA-FCAH

Kilian, P., Glaesser, J., Loose, F., & Kelava, A. (2021). Structure of pedagogical content knowledge in maths teacher education. *Psychological Test and Assessment Modeling, 63*, 337–360.

Knezek, G., & Christensen, R. (2016). Extending the will, skill, tool model of technology integration: Adding pedagogy as a new model construct | SpringerLink. *Journal of Computing in Higher Education, 28*, 307–325. https://doi.org/10.1007/s12528-016-9120-2

Koehler, M., Seob, S., & Mishra, P. (2011). How do we measure TPACK? Let me count the ways. In *Educational technology, teacher knowledge, and classroom impact: A research handbook on frameworks and approaches* (pp. 16–31). https://doi.org/10.4018/978-1-60960-750-0.ch002

Koh, J. H. L., & Chai, C. S. (2011). *Modeling pre-service teachers' technological pedagogical content knowledge (TPACK) perceptions: The influence of demographic factors and TPACK constructs*. December https://repository.nie.edu.sg/handle/10497/14126.

Koh, J. H. L., Chai, C. S., & Tsai, C. C. (2010). Examining the technological pedagogical content knowledge of Singapore pre-service teachers with a large-scale survey. *Journal of Computer Assisted Learning, 26*(6), 563–573. https://doi.org/10.1111/j.1365-2729.2010.00372.x

Koh, J. H. L., Chai, C. S., & Tsai, C.-C. (2013). Examining practicing teachers' perceptions of technological pedagogical content knowledge (TPACK) pathways: A structural equation modeling approach. *Instructional Science, 41*(4), 793–809. https://doi.org/10.1007/s11251-012-9249-y

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*, 716–725. https://doi.org/10.1037/0022-0663.100.3.716

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*(6), 1121–1134. https://doi.org/10.1037/0022-3514.77.6.1121

Lachner, A., Backfisch, I., & Franke, U. (2024). Towards an integrated perspective of teachers' technology integration: A preliminary model and future research directions. *Frontline Learning Research, 12*(1). https://doi.org/10.14786/flr.v12i1.1179. Article 1.

Lachner, A., Backfisch, I., & Stürmer, K. (2019). A test-based approach of modeling and measuring technological pedagogical knowledge. *Computers & Education*. https://doi.org/10.1016/j.compedu.2019.103645

Lachner, A., Fabian, A., Franke, U., Preiß, J., Jacob, L., Führer, C., Küchler, U., Paravicini, W., Randler, C., & Thomas, P. (2021). Fostering pre-service teachers' technological pedagogical content knowledge (TPACK): A quasi-experimental field study. *Computers & Education, 174*, Article 104304. https://doi.org/10.1016/j.compedu.2021.104304

Maderick, J. A., Zhang, S., Hartley, K., & Marchand, G. (2016). Preservice teachers and self-assessing digital competence. *Journal of Educational Computing Research, 54*(3), 326–351. https://doi.org/10.1177/0735633115620432

Maki, R. H., & McGuire, M. J. (2002). Metacognition for text: Findings and implications for education. In T. J. Perfect, & B. L. Schwartz (Eds.), *Applied metacognition* (1st ed., pp. 39–67). Cambridge University Press. https://doi.org/10.1017/CBO9780511489976.004.

Max, A., Lukas, S., & Weitzel, H. (2022). The relationship between self-assessment and performance in learning TPACK: Are self-assessments a good way to support preservice teachers' learning? *Journal of Computer Assisted Learning, 38*(4), 1160–1172. https://doi.org/10.1111/jcal.12674

Mishra, P. (2019). Considering contextual knowledge: The TPACK diagram gets an upgrade. *Journal of Digital Learning in Teacher Education, 35*(2), 76–78. https://doi.org/10.1080/21532974.2019.1588611

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record: The Voice of Scholarship in Education, 108*(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

Pamuk, S., Ergun, M., Cakir, R., Yilmaz, H. B., & Ayas, C. (2015). Exploring relationships among TPACK components and development of the TPACK instrument. *Education and Information Technologies, 20*(2), 241–263. https://doi.org/10.1007/s10639-013-9278-4

Sailer, M., Stadler, M., Schultz-Pernice, F., Franke, U., Schöffmann, C., Paniotova, V., Husagic, L., & Fischer, F. (2021). Technology-related teaching skills and attitudes: Validation of a scenario-based self-assessment instrument for teachers. *Computers in Human Behavior, 115*, Article 106625. https://doi.org/10.1016/j.chb.2020.106625

Saks, K., & Leijen, A. (2019). Digital learning diary as a tool for enhancing EFL learners' metacognitive reflection. In *2019 IEEE 19th international conference on advanced learning technologies (ICALT)* (pp. 263–264). https://doi.org/10.1109/ICALT.2019.00085

Scherer, R., Tondeur, J., & Siddiq, F. (2017). On the quest for validity: Testing the factor structure and measurement invariance of the technology-dimensions in the Technological, Pedagogical, and Content Knowledge (TPACK) model. *Computers & Education, 112*, 1–17. https://doi.org/10.1016/j.compedu.2017.04.012

Schmid, M., Brianza, E., Mok, S. Y., & Petko, D. (2024). Running in circles: A systematic review of reviews on technological pedagogical content knowledge (TPACK). *Computers & Education, 105024*. https://doi.org/10.1016/j.compedu.2024.105024

Schmid, M., Brianza, E., & Petko, D. (2020). Developing a short assessment instrument for Technological Pedagogical Content Knowledge (TPACK.xs) and comparing the factor structure of an integrative and a transformative model. *Computers & Education, 157*, Article 103967. https://doi.org/10.1016/j.compedu.2020.103967

Schmidt, D., Baran, E., Thompson, A., Mishra, P., Koehler, M., & Seob, S. (2009). Technological pedagogical content knowledge (TPACK): The development and validation of an assessment instrument for preservice teachers. *Journal of Research on Technology in Education, 42*, 123–149. https://doi.org/10.1080/15391523.2009.10782544

Schmidt-Atzert, L., & Amelang, M. (2012). *Psychologische diagnostik* (5. *Vollständig überarbeitete und erweiterte Aufl)*. Springer.

Schraw, G. (2009). A conceptual analysis of five measures of metacognitive monitoring. *Metacognition and Learning, 4*(1), 33–45. https://doi.org/10.1007/s11409-008-9031-3

Schubatzky, T., Burde, J.-P., Große-Heilmann, R., Haagen-Schützenhöfer, C., Riese, J., & Weiler, D. (2023). Predicting the development of digital media PCK/TPACK: The role of PCK, motivation to use digital media, interest in and previous experience with digital media. *Computers & Education, 104900*. https://doi.org/10.1016/j.compedu.2023.104900

Shrout, P. E., & Yip-Bannicq, M. (2017). Inferences about competing measures based on patterns of binary significance tests are questionable. *Psychological Methods, 22*(1), 84–93. https://doi.org/10.1037/met0000109

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4–14. https://doi.org/10.2307/1175860

Tondeur, J., van Braak, J., Guoyuan, S., Voogt, J., Fisser, P., & Ottenbreit-Leftwich, A. (2012). Preparing pre-service teachers to integrate technology in education: A synthesis of qualitative evidence. *Computers & Education, 59*(1), 131–144. https://doi.org/10.1016/j.compedu.2011.10.009

Urhahne, D., & Wijnia, L. (2021). A review on the accuracy of teacher judgments. *Educational Research Review, 32*, Article 100374. https://doi.org/10.1016/j.edurev.2020.100374

von Kotzebue, L. (2022). Two is better than one—examining biology-specific TPACK and its T-dimensions from two angles. *Journal of Research on Technology in Education, 0*(0), 1–18. https://doi.org/10.1080/15391523.2022.2030268

Voogt, J., Fisser, P., Pareja Roblin, N., Tondeur, J., & van Braak, J. (2013). Technological pedagogical content knowledge - a review of the literature: Technological pedagogical content. *Journal of Computer Assisted Learning, 29*(2), 109–121. https://doi.org/10.1111/j.1365-2729.2012.00487.x

Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika, 54*, 427–450. https://doi.org/10.1007/BF02294627

Willermark, S. (2018). Technological pedagogical and content knowledge: A review of empirical studies published from 2011 to 2016. *Journal of Educational Computing Research, 56*(3), 315–343. https://doi.org/10.1177/0735633117713114