Computers and Education Open 5 (2023) 100152

Contents lists available at ScienceDirect

Computers and Education Open

journal homepage: www.sciencedirect.com/journal/computers-and-education-open

Development and validation of a test to assess teachers' knowledge of how to operate technology

Tim Fütterer [a,*], Ronja Steinhauser [b], Steffen Zitzmann [a], Katharina Scheiter [c], Andreas Lachner [a], Kathleen Stürmer [a]

[a] Hector Research Institute of Education Sciences and Psychology, University of Tübingen, Europastraße 6, Tübingen 72072, Germany
[b] University of Mannheim, Mannheim, Germany
[c] University of Potsdam, Potsdam, Germany

ARTICLE INFO

Keywords:
Technological knowledge
Teaching with technology
Teachers
Assessment
TPACK

ABSTRACT

To effectively adopt technology during teaching, teachers require knowledge of how to operate technology. Especially first-time technology users need knowledge of how to handle digital devices and software programs as a foundation to use technology in the classroom successfully. This knowledge has so far been assessed mainly using self-reports. However, self-assessments are insufficient for assessing knowledge as their validity is limited. Moreover, the few tests that exist to measure technological knowledge (TK) show weaknesses (e.g., lack of ecological validity, outdated items). We present a test assessing teachers' TK that is independent of specific operating systems, covers technology that is relevant in everyday teaching, and is grounded in acknowledged psychometric modeling principles. We iteratively developed a test (named T-TK) comprising 26 items, utilizing cognitive testing, expert feedback, and two studies ($N_{pilot\ study} = 268$ pre-service and in-service teachers, $N_{main\ study} = 233$ in-service teachers) to filter items that did not match in content and were not Rasch conform. T-TK showed a satisfactory Andrich's reliability ($Rel_{\text{Andrich}} = 0.73$). Using the sample $N_{main\ study}$, correlations between T-TK and technological knowledge (self-report, $r = 0.52$), pedagogical knowledge (test scores, $r = 0.18$), and technological pedagogical knowledge (self-report, $r = 0.33$; test scores, $r = 0.46$) indicated convergent and discriminant validity. Thus, the T-TK proves to be a reliable and valid instrument to capture teachers' TK. The T-TK can be used both by practitioners not requiring any statistical knowledge (e.g., for individual diagnostics) and in research (e.g., to analyze teachers' TK).

Introduction

Integrating technology (e.g., digital devices or software programs) in classrooms is an important goal of education systems in the 21st century [1–3]. For instance, traditional analog media (e.g., books, blackboards) are increasingly being supplemented by digital media (e.g., tablet computers; [4]). In educational contexts, technology offers numerous potentials to enrich teaching and learning (e.g., creating new ways of presenting and handling information and embedding learning in different social contexts).

To exploit these potentials for successful student learning, teachers must integrate technology in a way that enhances teaching quality [5–7]. Therefore, teachers are referred to as "keystone species" ([8], p. 439) or "essential agents" ([9], p. 111) when it comes to the use of technology in the classroom. However, to make high-quality integration of technology in the classroom more likely, an important prerequisite is that teachers possess professional knowledge like technological knowledge (TK; [10,11]). However, the successful use of technology in the classroom requires much more than TK (i.e., knowing how to use devices and applications; [12]), such professional knowledge is a fundamental prerequisite [13,14]. Indeed, TK showed a positive association with the frequency of using technology for teaching (e.g., [15,16]). Furthermore, TK is a prerequisite for developing technology-related skills like technological pedagogical skills [16,17]. Accordingly, Dong et al. [18], for instance, found that TK predicts technological pedagogical knowledge (TPK; e.g., creating new ways of presenting and handling information,

---

An ethics committee approved the study and the collection of the data and confirmed that the procedures were in line with ethical standards of research with human subjects (pilot study: date of approval: 28/01/2021, file number: AZ.: A2.5.4–151_ns; main study: date of approval: 03/12/2020, file number: AZ.: A2.5.4–148_ns).

[*] Corresponding author.

E-mail address: tim.fuetterer@uni-tuebingen.de (T. Fütterer).

https://doi.org/10.1016/j.caeo.2023.100152
Received 12 February 2023; Received in revised form 5 October 2023; Accepted 13 November 2023

embedding learning in different social contexts).

Teachers' TK has predominantly been measured using self-reports [19–21]. Self-reports are valuable for several reasons, for instance, because they can be used cost-efficiency and provide a reliable assessment of self-efficacy [19,22,23]. However, self-reports entail various disadvantages, such as their fallibility due to biases in evaluating one's abilities [21]. Objective tests that assess TK are scarce and show weaknesses. For instance, items often refer to outdated technology or are formulated depending on specific operating systems (e.g., Windows). The absence of tests limits the ability to study teachers' TK and its role in technology-enhanced teaching. Moreover, targeted support of teachers (e.g., by providing differentiated professional development [PD]) can be realized only to a limited extent because precise individual diagnostics are made difficult by the lack of objective instruments [24].

Therefore, in the present paper, we introduce a test (T-TK) to assess teachers' knowledge of how to use technology (as a basic facet of TK), aiming to overcome such known weaknesses of existing tests. The T-TK was specifically developed to reliably and validly capture teachers' professional (conceptional) knowledge of how to handle generic (e.g., word processing programs) and school-specific (e.g., open-source web tools) technologies that are relevant to everyday teaching across different subjects and operating systems. In particular, the test is designed to be independent of specific operating systems (e.g., Windows), manufacturers, tools, or applications, to be valid across school subjects (e.g., mathematics, English), and to satisfy psychometric properties (e.g., reliability).

Theoretical framework

The construct "Teachers' technological knowledge"

Nowadays, the use of technology in the classroom is becoming increasingly important. Thus, teachers must be equipped with professional knowledge to be able to use technology. The term (digital) *technology* encompasses hardware (physical features of computer technology such as a keyboard or memory unit) and software components that can be used with computer technology (e.g., word processing software; [25,26]). Teachers must possess both TK related to generic technology and TK related to school-specific technology. Generic technologies are digital devices and software programs used by a broad range of users and not explicitly developed for teaching (see also the definition of operational technology: [14]). Nevertheless, some of these technologies are regularly used for teaching (e.g., word processing programs; [2]). School-specific technologies are digital devices and software programs that have been specifically designed for teaching and learning (e.g., [2,27]; see also the definition of pedagogical technology: [14]) and used by teachers mainly in a professional context (e.g., tutoring software: [28]).

Although the successful use of technology in the classroom requires much more than TK, TK and especially professional knowledge of how to handle digital devices and software applications (a core aspect of TK) is a fundamental prerequisite [13,14,27]. In other words, if teachers do not possess knowledge of how to operate technology, pedagogical-psychological or subject-didactic considerations regarding the use of technology are invalid.

A prominent model that highlights the role of TK and conceptualizes the required professional knowledge facets is the *Technological Pedagogical Content Knowledge* (TPACK) framework [11]. In this model, Mishra and Koehler [11] added TK as a third basic knowledge facet beside the well-known knowledge facets *pedagogical* and *content knowledge* postulated by Shulman [29]. Koehler et al. [10] specified TK as "knowledge about traditional and new technologies" (p. 102) that can be used in the classroom. The TPACK framework distinguishes between general knowledge domains like *pedagogical knowledge* (PK) and technology-related knowledge facets like TK or the intersection of PK and TK labeled as *technological pedagogical knowledge* (TPK; [10,30]). All of these knowledge facets are considered important for the effective use of technology in the classroom [31,32] and TK is an important aspect in numerous frameworks of technological competencies (e.g., teacher digital competency framework: [13]; European framework for the digital competence of educators [DigCompEdu]: [33]; for an overview, see [9]). All these frameworks include technological knowledge as a central basic prerequisite for effective teaching with technology. For instance, Falloon [13] distinguishes between *technical* (i.e., knowledge of mechanical fundamentals of technology) and *technological* knowledge (i.e., knowledge about the role and potential of technology in teaching).

Due to the rapid development of technology [20,26] and ambiguous conceptualizations of TK [19,34,35], the definitions of TK are vague, malleable, and differ regarding the hardware and software covered by the term technology. Thus, after their initial TK definition, Koehler and Mishra [25] linked TK to the notion of Fluency of Information Technology (FITness; [36]). FITness requires "that persons understand information technology broadly enough to apply it productively at work and in their everyday lives, to recognize when information technology can assist or impede the achievement of a goal, and to continually adapt to changes in information technology" ([26], p. 15). This definition emphasizes that TK needs to develop in line with technological innovation. That is, TK is in a "state of flux" and evolves over a lifetime [26]. Accordingly, teachers must constantly learn how to deal with new technology. In line with this reasoning, Angeli and Valanides [37] defined TK (more precisely, ICT knowledge) as "knowing how to operate a computer and knowing how to use a multitude of tools/software as well as troubleshoot in problematic situations" (p. 158). However, as TK is often discussed as knowledge about a technology's affordances for learning and teaching, there is controversy over how TK can precisely be separated from other knowledge domains such as TPK [34].

The relation of technological knowledge to other knowledge facets

In general, findings indicate that teachers' technology-related knowledge facets constitute a unique component of teachers' professional knowledge [38]. As TPACK knowledge facets are predominantly measured with self-reports [10], it isn't easy to ascertain the relationships of TK to other knowledge domains. However, the distinction from other general knowledge facets such as PK is apparent in previous findings. For instance, in an online study with 120 participants (i.e., in-service teachers) Lachner et al. [39] found only a weak positive association between test-based TK and PK ($r = 0.27$). Schmidt et al. [38] also found a weak positive association between self-reported TK and PK ($r = 0.21$) in their study on pre-service teachers' TPACK.

Regarding TPK, findings from previous research show that TK is positively associated with TPK [18,39–42]. For instance, Schmidt et al. [38] found a moderate ($r = 0.40$) and Scherer et al. [[32]; $N = 688$ pre-service teachers] found a strong ($r = 0.82$) positive association between self-reported TK and TPK. In contrast, Lachner et al. [39] found weak associations (e.g., $r = 0.23$) between test-based TK and TPK. However, in this study TK was assessed using the Test of Technological and Information Literacy (TILT: [43]) which is a test that measures not specifically TK but the broader construct computer literacy.

Measuring teachers' technological knowledge

Teachers' TK has predominantly been measured using self-reports [20,21,35]. In these self-report measures, teachers are asked to report how well they can solve technical problems [38]; how much they know about using word processing, spreadsheet, and presentation programs [44]; or how well they can, for instance, operate an interactive whiteboard [45]. For instance, Howard et al. [46] used a self-efficacy scale where teachers self-assessed knowledge facets where technological knowledge is involved (e.g., TCK) their ability to conduct online teaching in secondary education (e.g., "My ability to use various courseware programs to deliver instruction (e.g., Blackboard, Centra).").

Using self-report instruments has numerous advantages, such as cost-efficiency, a reliable assessment of self-efficacy, or a good prediction of teachers intention to use technology for teaching [22,23,32,47].

However, the validity of self-report instruments is limited [19,20,23,48]. The accuracy of self-reports is affected by teachers' ability to assess their knowledge [21,49–51] and by social desirability [52,53]. Moreover, self-reports are opposed to reflecting true knowledge in a specific domain [22,23,32,39,54,55]. Alternative ways of measuring teachers' TK are design tasks (e.g., creating lesson plans; [56]) and teacher observations ([56,57]; for an instrument see: [58]). However, teachers' TK is difficult to observe (see [57]), and observations are time- and cost-intensive [56].

Against the backdrop of weaknesses of self-reports and observations, tests appear to be an objective and efficient way to measure teachers' TK. A few tests already exist that measure TK or TK-related constructs (e.g., [43,59–62]). For instance, the INCOBI-R measures the theoretical and practical computer knowledge (computer literacy) of the general population [60] and the Basic Computer Proficiency (BCP) test covers various computer-related topics such as e-mail, computer accessories (e.g., printers), and databases [59]. Senkbeil and Ihme ([43]; see also [63]) developed the TILT measuring technology and information literacy by assessing conceptual knowledge about software applications such as Microsoft Word or Excel. This test is convincingly developed both theoretically and psychometrically but measures TK only as part of a broader construct of computer literacy and not exclusively.

In general, however, existing tests show some weaknesses that limit their use in assessing teachers' TK. First, most of existing tests were not developed for the target group of teachers, thereby lacking reference to technology relevant to teachers (e.g., learning management systems). However, the inclusion of such technology is important to ensure the tests' ecological validity for everyday teaching with technology. Second, some tests do not specifically capture TK but a broader construct such as ICT literacy that includes, but is not limited to TK. Third, existing tests are often focused on a specific operating system (e.g., Windows). As about a quarter of computer users work with other operating systems, such as macOS or Linux [64], tests focusing on a specific operating system discriminate against users of other operating systems. Fourth, due to the state of flux of TK some tests utilize outdated items (e.g., knowledge about defragmentation). This means that, in line with the ongoing development of technology, new test items are always needed to obtain reliable and valid statements on levels of TK.

Objectives and assumptions of the present paper

According to the recommendation for the development of a broader range of valid and reliable instruments to assess TPACK, such as standardized tests [65], the aim of this paper is to introduce a validated test of teachers' knowledge of how to operate technology (T-TK) as a core facet of TK. The T-TK was developed to meet requirements according to the limitations of the existing tests mentioned above. First, T-TK should assess teachers' conceptual knowledge of how to operate technology. Conceptual knowledge is defined as "static knowledge about facts, concepts, and principles" [[66], p. 107] that are used to solve problems and that support procedural knowledge [67]. Second, T-TK should be independent of specific operating systems (e.g., Windows), manufacturers, tools, or applications, to ensure that teachers who use different operating systems or applications are not discriminated against by the wording of the items. Third, T-TK should encompass technology that is currently relevant in everyday teaching. Furthermore, we aimed for T-TK to be valid across school subjects (e.g., mathematics, English) and to satisfy psychometric properties (e.g., reliability).

Regarding the test's validation, four assumptions based on previous research were investigated. First, based on a moderate positive association ($r = 0.46$) between test-based TK and self-reported knowledge of how to operate and apply technology [19] and as correlations between test scores and self-reports in general are often weak to moderate [51], we expected moderate positive correlations between T-TK and self-reported TK (*convergent validity*; assumption 1). Second, moderate positive correlations between self-reported TK and self-reported technological pedagogical knowledge (TPK) were found in previous studies (e.g., [16,38]). However, due to weaker correlations between test scores and self-reports in general, we expected weak positive correlations between T-TK and self-reported TPK (assumption 2). Third, based on the synopsis of the findings by Schmidt et al. [38], Scherer et al. [32], and Lachner et al. [39], we expected moderate positive correlations between T-TK and test-based TPK (assumption 3). Fourth, as TK and pedagogical knowledge (PK) are general knowledge facets that can be separated from each other, and based on the weak positive correlation between test-based TK and test-based PK reported by Lachner et al. [39], we expected a weak positive correlation between T-TK and test-based PK (*discriminant validity*; assumption 4).

Method

As we understand knowledge of how to operate technology as a unidimensional facet of TK with different content domains, our goal was to develop a test (T-TK) whose psychometric properties correspond to a unidimensional Rasch model. The advantage of a unidimensional model is that the sum score (i.e., the sum across the raw scores) is a sufficient statistic for the measured ability [68,69]. The sum score can then be easily used by practitioners who have little prior statistical experience. Further information on the Rasch model and its advantages can be found in Appendix A.

Development of the test T-TK

The development of the test was divided into two phases. In Phase 1, a theoretical concept was developed, and items were generated accordingly as well as revised after feedback that we received in a group discussion with experts. In Phase 2, preselected items were piloted.

Phase 1: theoretical conception and item creation

We defined knowledge of how to operate technology as a unidimensional facet of TK that encompasses different content areas. We stated that teachers need knowledge of both generic and school-specific technologies that are important in everyday teaching (e.g., [2]). For generic technology, we included knowledge of word processing, presentation, spreadsheet, email, and image and video editing programs as well as web browsers and digital devices. We categorized school-specific technology according to core teaching activities (i.e., "practices that occur with high frequency in teaching" [70], p. 277). Based on teaching activities related to the use of technology named in previous studies (e.g., [33,71]), we adapted four categories of teaching activities related to technology (Table 1).

However, school-specific technology can be further differentiated into technology that was not (e.g., spreadsheets) or was primarily developed to promote learning processes (e.g., learning apps, digital textbooks; [14]). The latter technology was not considered as this technology often requires little knowledge of how to operate it by the teacher.

To capture conceptual knowledge of teachers on how to operate technology used in teaching situations, we based the item formulation on situational judgment tests [72], in which people are asked to make judgments regarding situations encountered in the work place [19,73]. 138 single-choice items were generated based on the theoretical conceptualization of TK and especially knowledge of how to operate technology. Each item consisted of a task (item stem) and four possible solutions, one of which is correct. A direct reference to teaching was provided in the item stem. User manuals and the functionality of the technology were considered to ensure that all items can be answered without knowledge of specific technology.

The items were developed in several feedback loops in exchange with

Table 1
Core categories of teaching activities related to technology.

	Presenting and sharing information	Organizing learning processes	Regulating learning processes	Learning and content management systems
Description	All technology used to *display or transmit information.*	All technology used to *support the organization of learning processes.*	All technology used to *support self-regulated learning* (e.g., assessment, feedback, adaptive teaching).	All technology used to *organize digital distance teaching and learning* like internet-based courses.
Examples	screencasts, clouds	*knowledge management* (e.g., Etherpads, mindmap tools), *collaboration* (e.g., chats), *interactivity* (e.g., video conferencing tools)	audience response systems, quizzes	learning management systems

Note. Technology used for teaching can be categorized as technology that either was not (e.g., spreadsheets) or was primarily developed to promote learning processes (e.g., learning apps, digital textbooks). Technology that is already didactically designed and, thus, often requires little knowledge of how to operate it by the teacher, was not considered.

Table 2
Example Items of the T-TK Test.

Item wording	Response options
Generic Technology	
How do you go about emailing ten of your students without them seeing each other's email addresses?	I compose the email, then put the ten students in the Bcc, and finally send the email.
	I first create a distribution list with the email addresses of the ten students, then compose the email, and finally send the email to me and as a copy to the email addresses in the distribution list.
	I compose the email, then put the ten students in the cc, and finally send the email.
	I compose the email, then put the email addresses of the ten students into ten different address lines, and finally send the email.
School-Specific Technology	
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?	I can activate the function that the author colors are visible.
	I can activate the function that the students' name abbreviations are displayed before each of their entries.
	I can activate the function that only one student writes on the pad at a time.
	I can activate the function that each student writes on his or her own pad.

Note. The correct response options are printed in bold type.

experts and teachers.

Phase 1: expert discussion

We invited leading national experts with many years of experience in research on technology-enhanced teaching and learning and who know the German school system in detail to review all items critically. The experts received all items and the underlying theoretical conceptualization as well as the objectives of the test to be developed. They were asked both to give general feedback (e.g., item wording) and to provide a reasoned decision about whether an item should be excluded, included, or included but modified considering the goals of the test and the theoretical construct TK.

In a first step, six experts reviewed the item formulations and made their decisions individually.

In a second step, all experts discussed the items intensively over several hours in a focus group setting [74]. The focus group procedure is a highly structured survey method to obtain both individual and group statements on the items and thus to collect as broad a range of opinions as possible. The discussion reaffirmed theoretical understanding of a multidimensional construct of TK. Weaknesses of several items became apparent, which did not allow a clear delineation between the different facets of TK. In addition, the experts provided a lot of feedback about item content and formulations.

Based on the feedback of the experts, 53 items were preselected after a critical review of the content (i.e., fit to the construct of knowledge of how to operate technology) and adjusted in content and item formulation by the first and the second authors with advice from the co-authors.

Afterwards, the items were checked for content accuracy by a computer scientist. An overview of all items can be found in the online supplemental material (Open Science Framework [OSF] project: [75]), example items are shown in Table 2.

Phase 1: cognitive pretesting

Cognitive pretesting was applied to the 53 items [76]. To this end, four teachers were given the 53 items with a request to assess comprehensibility, confidence in answering, and relevance of the questioned topic for teaching for each item. In addition, teachers were given the opportunity to provide feedback on each item. All items were kept because none of the items showed alarming ratings by the teachers. However, small changes in wording were made due to these ratings.

Phase 2: pilot study

In Phase 2, a pilot study (60-minute online survey) was conducted to identify items with suboptimal fit to the Rasch model and to get a first impression of the distribution of item difficulties as well as the reliability of the status of the test. For the pilot study a sample of $N = 268$ pre-service and in-service teachers were recruited. On average, the participants were $M = 26.56$ ($SD = 7.99$) years old (77 % female).

Statistical analyses were conducted using *R* 4.0.4 (R [77]) and the package *eRm* [78]. Missing item answers were defined as missing values as this approach delivers almost unbiased results in competence tests [79]. After specifying a Rasch model with all 53 items, we first estimated item difficulties using *Conditional Maximum Likelihood* estimation and then estimated person parameters using *Maximum Likelihood*. Two item fit statistics were inspected: (1) the *Unweighted-Mean-Square-Statistic* (MNSQ; outfit) and (2) the *Weighted-Mean-Square-Statistic* (WMNSQ; infit; [80]). As a rule of thumb, infit and outfit values between 0.7 and 1.3 are acceptable [81]. Due to economic reasons (i.e., acceptable test length for use in research projects), in the pilot study, we chose even stricter criteria. That is, values below 0.7 were defined as indicating overfit, whereas values above 1.1 indicated underfit.

Overfitting indicates that the item fits the model better than expected by the Rasch model (less variation/noise than expected), whereas underfitting indicates too little item-model fit (more variation/noise than expected; [82]). Items that were above or below the thresholds of infit or outfit were excluded. After excluding 17 items, 36 items remained for use in the subsequent main study (for detailed information of the selection process, see the online supplemental material).

Phase 2: main study

A second study (60-min online survey) was conducted to test the items further. For the main study, data from $N = 233$ academic track (Gymnasium) teachers were collected between May and June 2021 as part of a PD course that focused on effective use of technology in the classroom. On average, the participants were $M = 43.25$ ($SD = 9.24$) years old (75 % female) and had an average of $M = 12.94$ years of teaching experience ($SD = 8.12$).

Participants were given 37 items, the 36 resulting items from the pilot study and one additional item, which we decided to include for content reasons. The statistical analyses and procedures were the same as those used in the pilot study. For detailed information of statistical properties (e.g., ICCs) and the item selection process, see the online supplemental material. We selected 26 items for the final T-TK (Table A in the appendix). Twenty-one items cover generic technology, and 5 items cover school-specific technology. Items covering the topics *presenting and sharing information* and *regulating learning processes* were eliminated and were thus no longer part of the final test (see also Discussion). The unit scale of the T-TK to assess teachers' knowledge of how to operate technology is in logits.

Validation of the test

Sample

The sample of the main study was used to validate the T-TK.

Measures

To validate the T-TK, we used self-reported TK and TPK as well as test-based TPK and PK. In particular, the associations between T-TK and self-reported TK to test convergent validity and T-TK and test-based PK to test discriminant validity are of central importance.

Self-reported technological knowledge. Teachers' self-reported technological knowledge (TK) was assessed with seven items (e.g., "I know how to solve my own technical problems") of the TPACK survey by Schmidt et al. [38]. The items were rated on a 5-point rating scale (1 = *strongly disagree* to 5 = *strongly agree*). We computed a manifest scale score. In our data, Cronbach's Alpha (α) was 0.95, indicating an excellent internal consistency [83].

Self-reported technological pedagogical knowledge. Teachers' self-reported TPK was assessed with five items (e.g., "I can choose technologies that enhance the teaching approaches for a lesson") of the TPACK survey by Schmidt et al. [38]. The items were rated on a 5-point rating scale (1 = *strongly disagree* to 5 = *strongly agree*). We computed a manifest scale score. In our data, internal consistency was satisfactory ($\alpha = 0.71$; [83]).

Test-based technological pedagogical knowledge. A test was used to assess teachers' technological pedagogical knowledge (TPK) consisting of eight open-ended questions where teachers are confronted with different teaching situations (Table 3; Franke et al. [84]). For each teaching situation, teachers were asked to answer whether and how technology can be used in a didactically meaningful way.

In total, a maximum of 3 points could be achieved for each task. Points were awarded based on the relation to the quality of instruction (e.g., cognitive activation); whether a teacher uses technology as a replacement, amplification, or transformation (RAT framework; [85]) and the quality of the justification of the answer. Responses were scored by two extensively trained raters based on a coding manual. Tests from 40 participants (17 %) were rated by both raters. A weighted kappa of $\kappa = 0.81$ ($p < .001$) indicated a satisfactory inter-rater reliability [86]. As recommended by Franke et al. [84], the sum score was used in the analyses of this study.

Table 3
Example item of the technological pedagogical knowledge test.

Item wording	Question
In the previous lesson, your students learned a new basic skill. As a teacher, you want your students to continue practicing and consolidating this skill as homework.	How could you use educational technologies to support the students' practicing? Please give reasons for your answer.

Note. An open-ended response format was used.

Test-based pedagogical knowledge. Teachers' pedagogical knowledge (PK) was assessed using the short-scale *teaching* of the PK test [87] consisting of 15 single- and multiple-choice items with four response options each (Table 4).

This short scale has proven to be reliable ($0.65 < \text{EAP/PV} < 0.76$: [87]). Two points were awarded for each correct single-choice task. For multiple-choice tasks, points were awarded as follows: 0 points for 0 or 1 correct response, 1 point for 2 or 3 correct responses, and 2 points for 4 correct responses. As recommended by Kunter et al. [87], the sum score was used in the analyses of this study.

Statistical analyses

For the subsequent analyses, the R package *lavaan* [88] was used.

We employed a structural equation model (SEM) approach to test whether the data fit the Rasch model. As recommended for models with dichotomous items [89], estimation was carried out with robust weighted least squares (WLSMV).

Andrich's reliability was chosen as the reliability estimate of the test [90]. This reliability can be interpreted similarly to Cronbach's α [91].

To investigate the discriminant and convergent validity of the T-TK, correlations between the participants' abilities (i.e., person parameters from the Rasch model) and constructs of interest (i.e., self-reported TK, self-reported and test-based TPK, and test-based PK) were calculated. As Spearman rank-order correlations are known to be robust against violations of the normal distribution assumption [92], they were computed to test associations between the constructs. To evaluate the statistical significance of the correlations, type 1 error rate was fixed at 5 % (two-sided).

There were only a few missing values on the variables of interest (0.5 %). Therefore, as recommended by J.W. Graham et al. [93], casewise deletion was applied in order to treat missing values.

Results

As can be seen in Table 5, the mean value of test-based TPK was rather low ($M = 7.84$ out of a total possible score of 24), whereas the mean values of the other scales were in the expected range.

For the 26 final test items of the T-TK, a Rasch model was specified. Means, standard deviations, item difficulties, and infit and outfit statistics of the items are presented in Table 6.

Mean values were between $M_{TK14/TK18} = 0.50$ and $M_{TK4} = 0.94$, and item difficulties were between $\sigma_{TK4} = -2.11$ and $\sigma_{TK14} = 1.08$. The ICCs for all 26 items are shown in Fig. 1.

Table 4
Example Item of the Pedagogical Knowledge Test.

Item wording	Subquestion
A positive learning climate is created …	through the interplay of students' experience of autonomy, social inclusion and experience of competence in the classroom. when teaching is determined by mutual respect, adherence to rules, shared responsibility, justice and caring. when project work is used more frequently in the classroom. when students can choose materials independently.

Note. Correct response options are printed in bold type.

Table 5

Descriptive statistics for the measures used for validation.

	N	*M*	*SD*	*Min*	*Max*
T-TK test[a]	232	1.12	1.12	−1.66	4.33
TK (self-reported)	233	3.32	1.04	1	5
TPK (self-reported)	233	3.31	0.65	1	5
TPK test	230	7.84	2.93	0.5	14
PK test	233	19.89	2.82	12	27

Note. T-TK = teachers' knowledge of how to operate technology; TK = technological knowledge; TPK = technological pedagogical knowledge; PK = pedagogical knowledge.

[a] The unit scale is in logits.

Table 6

Means, standard deviations, item difficulties, and infit and outfit statistics for T-TK.

Item name	*n*	*M*	*SD*	σ	Infit	Outfit
TK1	232	0.88	0.32	−1.25	1.00	0.98
TK2	231	0.54	0.50	0.87	1.07	1.09
TK3	232	0.89	0.31	−1.35	0.90	0.63
TK4	233	0.94	0.23	−2.11	0.96	0.84
TK5	227	0.76	0.43	−0.30	0.98	0.94
TK6	226	0.73	0.44	−0.13	0.99	1.02
TK7	230	0.88	0.33	−1.21	0.98	1.15
TK8	229	0.81	0.39	−0.59	0.91	0.90
TK9	226	0.68	0.47	0.21	1.00	0.93
TK10	222	0.55	0.50	0.84	0.95	0.91
TK11	221	0.69	0.46	0.12	0.91	0.87
TK12	231	0.83	0.38	−0.75	0.98	0.91
TK13	231	0.88	0.32	−1.25	0.86	0.97
TK14	227	0.50	0.50	1.08	1.03	1.03
TK15	222	0.64	0.48	0.38	1.11	1.15
TK16	220	0.57	0.50	0.75	0.99	0.97
TK17	219	0.53	0.50	0.90	1.06	1.07
TK18	228	0.50	0.50	1.06	0.82	0.77
TK19	228	0.65	0.48	0.33	1.02	1.03
TK20	224	0.75	0.43	−0.24	1.04	1.02
TK21	224	0.71	0.45	0.01	1.09	1.13
TK22	220	0.56	0.50	0.80	1.11	1.22
TK23	217	0.52	0.50	0.99	0.90	0.91
TK24	223	0.53	0.50	0.93	1.05	1.06
TK25	225	0.68	0.47	0.22	0.90	0.87
TK26	220	0.76	0.43	−0.31	1.04	1.05

Note. σ = Item Difficulty; Infit = Weighted-Mean-Square-Statistic; Outfit = Unweighted-Mean-Square-Statistic.

A graphical representation of the distribution of item difficulties can be found in the Wright map (Fig. 2). In Fig. 2 the unit scale is in logits.

As can be seen, most items are approximately located between 0 and 1 logits. Some items can also be found between −2 and 0 logits. Few items are located above 1 logit. That is, the high ability range (> 1 logit) is barely covered by items. The majority of the surveyed teachers fall within the range of 0 to 2 logits in terms of their ability.

Following Bentler [94], comparative fit index (CFI) = 0.928, and following Hu and Bentler [95], root mean square error of approximation (RMSEA) = 0.027 (90 % CI [.009, 0.039]), indicated a satisfactory fit of the model, whereas the χ^2-goodness-of-fit-test (χ^2 [324, *N* = 233] = 373.25, *p* = .031) and the standardized root mean square residual (SRMR) = 0.131 indicated an unsatisfactory model fit. However, the χ^2-test has to be interpreted with caution as it is influenced by factors such as the sample size [96]. As most of the absolute and incremental fit indices (χ^2 excluded) indicated an acceptable model fit, the model fit was considered to be valid. Andrich's coefficient was $Rel_{\text{Andrich}} = 0.73$, which indicated acceptable reliability.

The T-TK correlated statistically significantly with all other constructs assessed (Table 7). The Spearman rank-order correlations were calculated because the Shapiro-Wilk test indicated that the data were not normally distributed (T-TK test: *W*[233] = 0.98, *p* = .006; self-reported TK: *W*[233] = 0.97, *p* < .001; self-reported TPK: *W*[233] = 0.99, *p* = .033; TPK test: *W*[233] = 0.99, *p* = .054; PK test: *W*[233] = 0.98, *p* = .002).

As predicted in assumption 1, we found a moderate positive [97] statistically significant correlation between T-TK and self-reported TK (*r* = 0.52, *p* < .001), indicating convergent validity of the T-TK.

As predicted in assumption 2 and in line with theory stating that TK and TPK are related but distinct constructs [11], we found a weak positive statistically significant correlation between T-TK and self-reported TPK (*r* = 0.33, *p* < .001).

As predicted in assumption 3, we found a moderate positive statistically significant correlation between the T-TK and the TPK test (*r* = 0.46, *p* < .001).

Regarding assumption 4 and as expected, we found a weak positive statistically significant correlation between the T-TK and the PK test (*r* = 0.18, *p* = .005).

Discussion

This study contributes to the research on teachers' TK by presenting the test T-TK, which measures teachers' knowledge of how to operate technology. More specifically, T-TK captures teachers' professional knowledge of how to handle generic (e.g., word processing programs) and school-specific (e.g., open-source web tools) technology that is relevant in everyday teaching across subjects and operating systems (e.g., Windows).

Our findings suggested that the T-TK can measure teachers' knowledge of how to operate technology economically, reliably, and validly.

First, the moderate correlation of T-TK with self-reported TK indicated convergent validity. Bearing in mind that the self-report instrument by Schmidt et al. [38] captures TK in rather general terms (e.g., "I know about a lot of different technologies"), whereas the T-TK assesses knowledge about technologies that are relevant specifically in teaching, the strength of the correlation is remarkable. Moreover, this correlation is in line with current research on correlations of objective assessments of TK and self-reported TK (especially when focusing on operating technology; [19]). However, findings from Akyuz [98] indicate that a greater correlation between self-reports and tests is plausible in knowledge domains that do not include a pedagogy aspect (e.g., TK).

Second, the weak correlation with PK indicated discriminant validity. This correlation is even weaker than the one found by Lachner et al. [39]. That is, our correlation indicates further evidence of discriminant validity of the T-TK.

Third, the moderate correlation between the T-TK and the TPK test is largely in line with findings from previous research showing moderate correlations between the two constructs (e.g., [38]). On the other hand, Lachner et al. [39] found weaker correlations; however, it should be noted that these authors used the TILT [63] to capture teachers' TK, which was not specifically tailored to teachers and does not capture TK as precisely as the T-TK but rather captures a broader construct (i.e., ICT literacy). Furthermore, TK was assessed without any direct reference to teaching. Nevertheless, it can be debated whether these findings indicate validity of the T-TK.

Limitations

Despite the promising findings, some limitations should be mentioned.

First, the independence of the T-TK from specific technology (e.g., specific applications) and operating systems as one of the major strengths of the test is at the same time a weakness. That is, if the operations of different technologies (e.g., LibreOffice Calc, Apple Numbers, Microsoft Excel) were too different, we did not create items for these unique operations. Thus, outweighing relevance by uniformity in operation could have led to relevant content not being queried. For instance, learning management systems (LMS) differ greatly in their operations, and this led us to survey quite general operations, such as

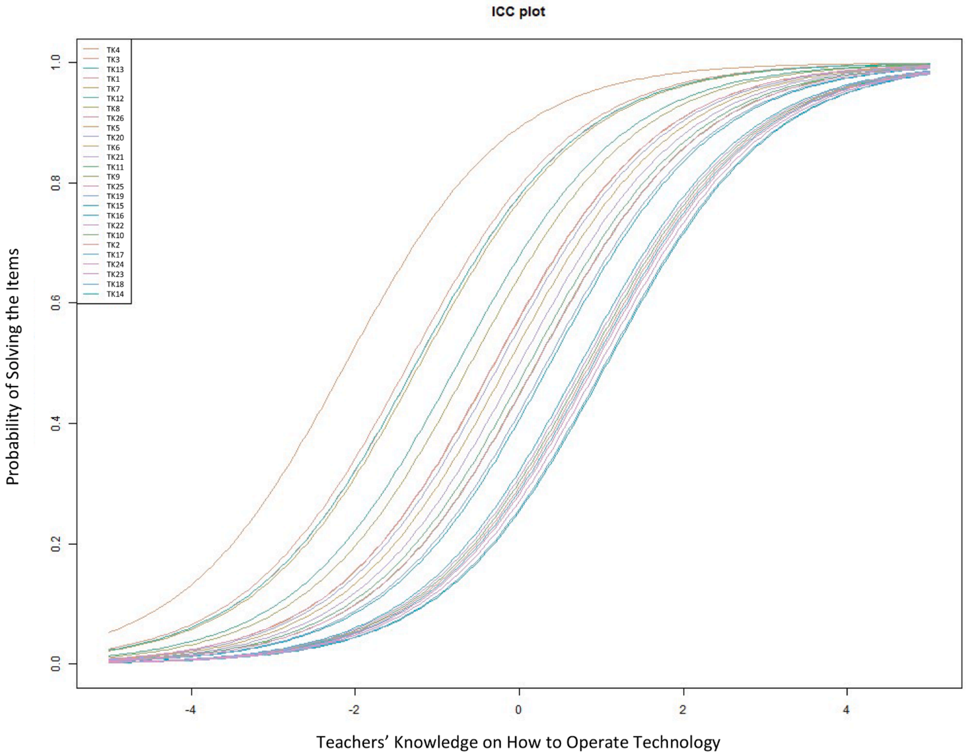

Fig. 1. *Item Characteristic Curves (ICC) for the Final Test T-TK Consisting of 26 Items. Note.* On the x-axis, the person's ability is mapped. On the y-axis, the probability of solving an item is depicted. The point at which the probability of solving an item correctly is 0.5 represents the item difficulty.

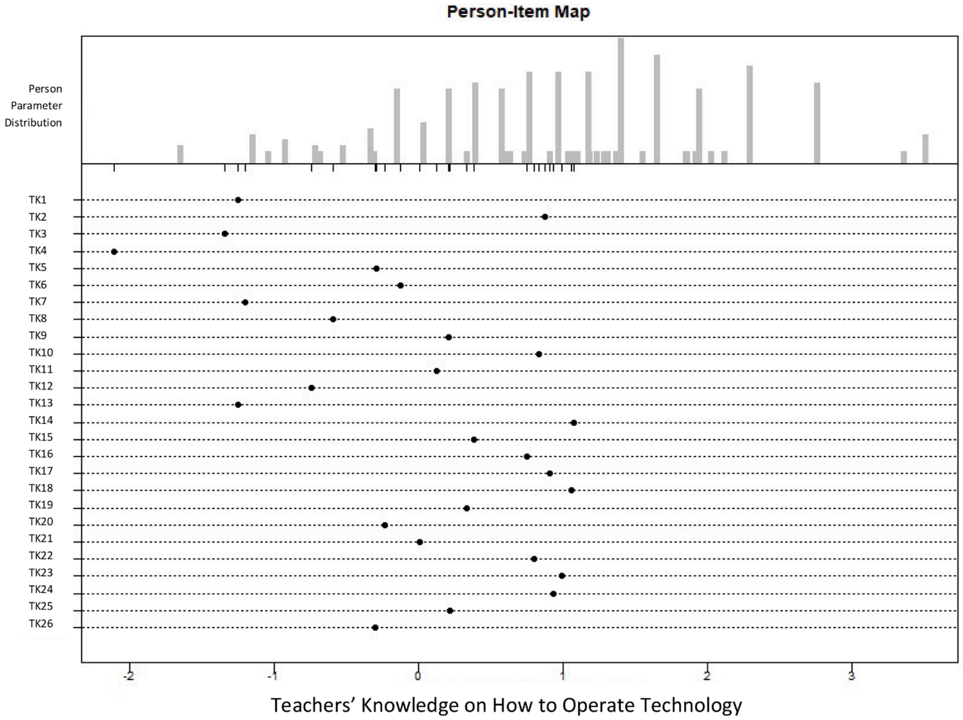

Fig. 2. *Wright Map of the Final Test T-TK Consisting of 26 Items. Note.* The unit scale is in logits. The columns in the upper part of the figure describe the distribution of participants' knowledge organized from lowest knowledge (left side) to highest knowledge (right side). The dots represent the item difficulties from easiest (left side) to most difficult (right side).

drag & drop. In other studies, however, specific LMS and their operations were surveyed (e.g., [99]). Furthermore, the items were presented exclusively in text format. In other studies, graphical supports are often provided (e.g., screenshots; [63]) that may make application scenarios more realistic. Nevertheless, the use of multiple-choice tasks seems to be appropriate for assessing TK [61].

Second, during the test development, many items covering school-specific technology had to be excluded due to insufficient fit to the

Table 7
Correlations among T-TK test, TK self-reports, TPK self-reports, PK test, and TPK test.

	(1)	(2)	(3)	(4)	(5)
(1) T-TK (test)	1				
(2) TK (self-reported)	.52, $p < .001$	1			
(3) TPK (self-reported)	.33, $p < .001$	.52, $p < .001$	1		
(4) TPK (test)	.46, $p < .001$	.30, $p < .001$	.27, $p < .001$	1	
(5) PK (test)	.18, $p = .005$	.02, $p = .705$	.17, $p = .011$	.23, $p < .001$	1

Note. T-TK = teachers' knowledge of how to operate technology; TK = technological knowledge; TPK = technological pedagogical knowledge; PK = pedagogical knowledge.

Rasch model. Items covering the topics *presenting and sharing information* and *regulating learning processes* were eliminated completely. Koehler et al.'s [10] definition of TK as the "knowledge about traditional and new technology" (p. 102) that can be used in the classroom actually implies that the T-TK should cover also a wide range of school-specific technology that is important for teaching and learning processes. However, a test that covers technology that could potentially be used in the classroom (e.g., feedback systems, collaborative software) may discriminate against those who do not use this technology in their teaching practice at all.

Implications and future research

Teachers' can significantly benefit from the T-TK because test results can be used for individualized diagnostics. Thus, appropriate supports can be provided to teachers on objectively identified needs, such as effective (e.g., adaptive and differentiated) professional development to match participants' ability [23,54]. Whereas teachers have often experienced a one-size-fits-all PD approach for technology integration in the past [26], tests like the T-TK allow PD providers to monitor teachers' levels of TK and to offer individualized (differentiated) PD [100]. That is, the simultaneous use of different tests measuring facets of professional knowledge could give insights into whether a teacher needs more support in technological or technological pedagogical topics.

With a scope of only 26 multiple-choice items, the T-TK can be used in a time-saving manner, and scoring is easily done by creating sum scores instead of logits. However, to enable even better economical usability of the test in practice, future research could develop a short version or a computer-adaptive version of the test.

Unlike previous research on teachers' TK, which has mainly relied on self-reports, future research with the T-TK allows one to investigate a variety of research questions test-based. First, it is thus possible to gain more robust insights into the role of TK regarding successful teaching with technology and its effect on teaching practices and students' learning. Second, the use of T-TK can also help to improve understanding of how TK evolves and what role TK plays in the acquisition of TPK compared to PK.

Regarding the focus of T-TK on knowledge of how to operate technology, which is on the one hand a core but on the other hand only one aspect of TK, and due to the fast development of technology, future research should further develop the T-TK.

First, items tapping into knowledge of the potentials and limitations of technology (e.g., data protection aspects) should be added, and items on outdated technology should be replaced by items on technology that may find their way into classrooms in the future (e.g., virtual reality).

Second, the high practical relevance and application proximity of the T-TK should be critically reviewed, especially with respect to school-specific technology.

Third, T-TK is intended to be usable by teachers of different school subjects, at different school types, and at different levels of TK. However, the final version of the T-TK was developed based on a sample of non-representative–academic track teachers in a PD context. The specific group of teachers could also be a reason for the moderate correlation between T-TK and self-reported TK as teachers in this study could have been more proficient in assessing their abilities, mainly because they may have reflected on their abilities before participating in the PD. Other studies are based on pre-service teachers (e.g., [40,101]) who are assumed to be less able to assess their abilities than in-service teachers due to a lack of experience [40]. Thus, the T-TK should be further validated, and more items of a high difficulty range should be included to better assess and differentiate teachers with higher TK.

In general, following Baier and Kunter [40] who advised using instruments other than self-reports to assess the validity of TK, future research could validate T-TK against other tests like the TILT for adults [43,63] or recently developed tests (e.g., [19]). Finally, future research should investigate whether teachers' knowledge of how to operate technology is indeed a unidimensional construct, as most items covering school-specific technology showed an insufficient fit to the Rasch model. This could be an indication that teachers' TK consists of two or even more facets: knowledge of generic technology, which is acquired in personal and professional contexts, and knowledge of school-specific technology, which is acquired exclusively in professional teaching contexts. Moreover, it is possible that a general cross-subject factor of knowledge of how to operate technology (i.e., generic applications) and other subject-dependent aspects of teachers' knowledge of how to operate technology (e.g., knowledge about technology in mathematics) exist. The latter idea is supported by studies that considered the use of technology in the classroom from a subject-specific viewpoint (e.g., [28, 98]).

To sum up, T-TK is the result of a first attempt to design a test that is meant to measure teachers' knowledge of how to operate technology as one facet of TK. On the one hand, T-TK is convincing by being Rasch conform, reliable, and valid, as well as by containing items that are independent of operating systems and that cover both generic and school-specific technology. On the other hand, T-TK should be further optimized in future studies and extended with respect to further facets of TK.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Data collection was supported by funds from the Dr. Hans Riegel Foundation, Bonn, Germany.

This research was funded by the Federal Ministry of Education and Research (BMBF) and the Baden-Württemberg Ministry of Science as part of the Excellence Strategy of the German federal and state governments.

This research was supported by the LEAD Graduate School & Research Network [GSC1028], which was funded within the framework of the Excellence Initiative of the German federal and state governments.

This research was supported by the Postdoctoral Academy of Education Sciences and Psychology of the Hector Research Institute of Education Sciences and Psychology, Tübingen, funded by the Baden-Württemberg Ministry of Science, Research, and the Arts.

This project is part of the 'Qualitätsoffensive Lehrerbildung', a joint initiative of the Federal Government and the Länder which aims to improve the quality of teacher training. The program is funded by the Federal Ministry of Education and Research.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.caeo.2023.100152.

References

[1] European Commission. Survey of schools. ICT in education: Benchmarking access, use and attitudes to technology in Europe's schools. Publications Office; 2013. https://doi.org/10.2759/94499.

[2] Fraillon J, Ainley J, Schulz W, Friedman T, Duckworth D. Preparing for life in a digital world: IEA International Computer and Information Literacy Study 2018. International report. Springer; 2020. https://doi.org/10.1007/978-3-030-38781-5.

[3] OECD. Skills for a digital world: Policy brief on the future of work. OECD Publishing; 2016. https://www.oecd.org/els/emp/Skills-for-a-Digital-World.pdf.

[4] Hammer M, Göllner R, Scheiter K, Fauth B, Stürmer K. For whom do tablets make a difference? Examining student profiles and perceptions of instruction with tablets. Comput. Educ. 2021;166:104147. https://doi.org/10.1016/j.compedu.2021.104147.

[5] Petko D, Cantieni A, Prasse D. Perceived quality of educational technology matters: A secondary analysis of students' ICT use, ICT-related attitudes, and PISA 2012 test scores. J. Educ. Comput. Res. 2017;54(8):1070–91. https://doi.org/10.1177/0735633116649373.

[6] Fütterer T, Scheiter K, Cheng X, Stürmer K. Quality beats frequency? Investigating students' effort in learning when introducing technology in classrooms. Contemp. Educ. Psychol. 2022;69:102042. https://doi.org/10.1016/j.cedpsych.2022.102042.

[7] Fütterer T, Hoch E, Lachner A, Scheiter K, Stürmer K. High-quality digital distance teaching during COVID-19 school closures: Does familiarity with technology matter? Comput. Educ. 2023;199:104788. https://doi.org/10.1016/j.compedu.2023.104788.

[8] Davis N, Eickelmann B, Zaka P. Restructuring of educational systems in the digital age from a co-evolutionary perspective: Restructuring of educational systems. J. Comput. Assist. Learn. 2013;29(5):438–50. https://doi.org/10.1111/jcal.12032.

[9] Almerich G, Orellana N, Suárez-Rodríguez J, Díaz-García I. Teachers' information and communication technology competences: A structural approach. Comput Educ 2016;100:110–25. https://doi.org/10.1016/j.compedu.2016.05.002.

[10] Koehler MJ, Mishra P, Kereluik K, Shin TS, Graham CR. The technological pedagogical content knowledge framework. In: Spector JM, Merrill MD, Elen J, Bishop MJ, editors. Eds..Handbook of research on educational communications and technology. Springer; 2014. p. 101–11. https://doi.org/10.1007/978-1-4614-3185-5_9.

[11] Mishra P, Koehler MJ. Technological pedagogical content knowledge: A framework for teacher knowledge. Teach Coll Rec 2006;108(6):1017–54. https://doi.org/10.1111/j.1467-9620.2006.00684.x.

[12] Janssen J, Stoyanov S, Ferrari A, Punie Y, Pannekeet K, Sloep P. Experts' views on digital competence: Commonalities and differences. Comput Educ 2013;68:473–81. https://doi.org/10.1016/j.compedu.2013.06.008.

[13] Falloon G. From digital literacy to digital competence: The teacher digital competency (TDC) framework. Educ Technol Res Dev 2020;68(5):2449–72. https://doi.org/10.1007/s11423-020-09767-4.

[14] Gao P.P., Nagel A., Biedermann H. Categorization of educational technologies as related to Pedagogical Practices. In: Tirri K., Toom A. Eds. Pedagogy in basic and higher education—current developments and challenges. IntechOpen; (2020). 10.5772/intechopen.88629.

[15] Cheng SL, Xie K. The relations among teacher value beliefs, personal characteristics, and TPACK in intervention and non-intervention settings. Teach Teach Educ 2018;74:98–113. https://doi.org/10.1016/j.tate.2018.04.014.

[16] Guggemos J, Seufert S. Teaching with and teaching about technology – evidence for professional development of in-service teachers. Comput Human Behav 2021;115:106613. https://doi.org/10.1016/j.chb.2020.106613.

[17] Pamuk S, Ergun M, Cakir R, Yilmaz HB, Ayas C. Exploring relationships among TPACK components and development of the TPACK instrument. Educ Inf Technol 2015;20(2):241–63. https://doi.org/10.1007/s10639-013-9278-4.

[18] Dong Y, Chai CS, Sang GY, Koh JHL, Tsai CC. Exploring the profiles and interplays of pre-service and in-service teachers' technological pedagogical content knowledge (TPACK) in China. Educ Technol Soc 2015;18(1):158–69. https://www.jstor.org/stable/jeductechsoci.18.1.158.

[19] Kastorff T, Sailer M, Vejvoda J, Schultz-Pernice F, Hartmann V, Hertl A, et al. Context-specificity to reduce bias in self-assessments: Comparing teachers' scenario-based self-assessment and objective assessment of technological knowledge. J Res Technol Educ 2022:1–14. https://doi.org/10.1080/15391523.2022.2062498.

[20] Koehler M.J., Shin T.S., Mishra P. How do we measure TPACK? Let me count the ways. In: Ronau R.N., Rakes C.R., Niess M.L., Eds. Educational technology, teacher knowledge, and classroom impact: A research handbook on frameworks and approaches. IGI Global; (2012). p. 16–31. 10.4018/978-1-60960-750-0.

[21] Willermark S. Technological pedagogical and content Knowledge: A review of empirical studies published from 2011 to 2016. J Educ Comput Res 2018;56(3):315–43. https://doi.org/10.1177/0735633117713114.

[22] Scherer R, Tondeur J, Siddiq F. 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. Comput Educ 2017;112:1–17. https://doi.org/10.1016/j.compedu.2017.04.012.

[23] Seufert S, Guggemos J, Sailer M. Technology-related knowledge, skills, and attitudes of pre- and in-service teachers: The current situation and emerging trends. Comput Human Behav 2021;115:106552. https://doi.org/10.1016/j.chb.2020.106552.

[24] Fütterer T, Scherer R, Scheiter K, Stürmer K, Lachner A. Will, skills, or conscientiousness: What predicts teachers' intentions to participate in technology-related professional development? Comput Educ 2023;198:104756. https://doi.org/10.1016/j.compedu.2023.104756.

[25] Koehler MJ, Mishra P. What is technological pedagogical content knowledge? Contemporary Issues in Technology and Teacher. Education 2009;9(1):60–70.

[26] Koehler MJ, Mishra P, Cain W. What is technological pedagogical content knowledge (TPACK)? Journal of Education 2013;193(3):13–9. https://doi.org/10.1177/002205741319300303.

[27] Horz H. Medien. In: Wild E., Möller J. Pädagogische Psychologie (3rd ed.) Springer; (2020). p. 133–160. 10.1007/978-3-662-61403-7.

[28] Hillmayr D, Ziernwald L, Reinhold F, Hofer SI, Reiss KM. The potential of digital tools to enhance mathematics and science learning in secondary schools: A context-specific meta-analysis. Comput Educ 2020;153:103897. https://doi.org/10.1016/j.compedu.2020.103897.

[29] Shulman LS. Those who understand: Knowledge growth in teaching. Educ Res 1986;15(2):4–14. https://doi.org/10.3102/0013189X015002004.

[30] Herring MC, Koehler MJ, Mishra P. Handbook of technological pedagogical content knowledge (TPACK) for educators. 2nd ed. Routledge; 2016. https://doi.org/10.4324/9781315771328.

[31] Koh JHL, Chai CS, Tay LY. TPACK-in-Action: Unpacking the contextual influences of teachers' construction of technological pedagogical content knowledge (TPACK). Comput Educ 2014;78:20–9. https://doi.org/10.1016/j.compedu.2014.04.022.

[32] Scherer R, Tondeur J, Siddiq F, Baran E. The importance of attitudes toward technology for pre-service teachers' technological, pedagogical, and content knowledge: Comparing structural equation modeling approaches. Comput Human Behav 2018;80:67–80. https://doi.org/10.1016/j.chb.2017.11.003.

[33] Redecker C, Punie Y. European framework for the digital competence of educators DigCompEdu. European framework for the digital competence of educators DigCompEdu; 2017.

[34] Graham CR. Theoretical considerations for understanding technological pedagogical content knowledge (TPACK). Comput Educ 2011;57(3):1953–60. https://doi.org/10.1016/j.compedu.2011.04.010.

[35] Voogt J, Fisser P, Pareja Roblin N, Tondeur J, van Braak J. Technological pedagogical content knowledge - a review of the literature. J Comput Assist Learn 2013;29(2):109–21. https://doi.org/10.1111/j.1365-2729.2012.00487.x.

[36] National Research Council. Being Fluent with Information Technology. The National Academies Press; 1999. https://doi.org/10.17226/6482.

[37] Angeli C, Valanides N. Epistemological and methodological issues for the conceptualization, development, and assessment of ICT–TPCK: Advances in technological pedagogical content knowledge (TPCK). Comput Educ 2009;52(1):154–68. https://doi.org/10.1016/j.compedu.2008.07.006.

[38] Schmidt DA, Baran E, Thompson AD, Mishra P, Koehler MJ, Shin TS. Technological pedagogical content knowledge (TPACK): The development and validation of an assessment instrument for pre-service teachers. J Res Technol Educ 2009;42(2):123–49. https://doi.org/10.1080/15391523.2009.10782544.

[39] Lachner A, Backfisch I, Stürmer K. A test-based approach of modeling and measuring technological pedagogical knowledge. Computers & Education 2019;142:103645. https://doi.org/10.1016/j.compedu.2019.103645.

[40] Baier F, Kunter M. Construction and validation of a test to assess (pre-service) teachers' technological pedagogical knowledge (TPK). Stud Educ Eval 2020;67:100936. https://doi.org/10.1016/j.stueduc.2020.100936.

[41] Celik I, Sahin I, Akturk AO. Analysis of the relations among the components of technological pedagogical and content knowledge (TPACK): A structural equation model. J Educ Comput Res 2014;51(1):1–22. https://doi.org/10.2190/EC.51.1.a.

[42] Koh JHL, Chai CS, Tsai CC. Demographic factors, TPACK constructs, and teachers' perceptions of constructivist-oriented TPACK. Educ Technol Soc 2014;17(1):185–96. https://www.jstor.org/stable/jeductechsoci.17.1.185.

[43] Senkbeil M, Ihme JM. NEPS technical report for computer literacy: Scaling results of starting cohort 6 - Adults, 61. Leibniz Institute for Educational Trajectories, National Educational Panel Study; 2015.

[44] Sahin I. Development of survey of technological pedagogical and content knowledge (TPACK). Turk Online J Educ Technol 2011;10(1):97–105.

[45] Jang SJ, Tsai MF. Exploring the TPACK of Taiwanese elementary mathematics and science teachers with respect to use of interactive whiteboards. Comput Educ 2012;59(2):327–38. https://doi.org/10.1016/j.compedu.2012.02.003.

[46] Howard S.K., Tondeur J., Siddiq F., Scherer R. Ready, set, go! Profiling teachers' readiness for online teaching in secondary education. Technology, Pedagogy and Education (2020);1–18. 10.1080/1475939X.2020.1839543.

[47] Aesaert K, Voogt J, Kuiper E, & van Braak J. Accuracy and bias of ICT self-efficacy: An empirical study into students' over- and underestimation of their ICT competences. Comput Human Behav 2017;75:92–102. https://doi.org/10.1016/j.chb.2017.05.010.

[48] Kopcha TJ, Ottenbreit-Leftwich A, Jung J, Baser D. Examining the TPACK framework through the convergent and discriminant validity of two measures. Comput Educ 2014;78:87–96. https://doi.org/10.1016/j.compedu.2014.05.003.

[49] Abbitt JT. Measuring technological pedagogical content knowledge in pre-service teacher education: A review of current methods and instruments. J Res Technol Educ 2011;43(4):281–300. https://doi.org/10.1080/15391523.2011.10782573.

[50] Drummond A, Sweeney T. Can an objective measure of technological pedagogical content knowledge (TPACK) supplement existing TPACK measures? Br J Educ Technol 2017;48(4):928–39. https://doi.org/10.1111/bjet.12473.

[51] Dunning D, Heath C, Suls JM. Flawed self-assessment: Implications for health, education, and the workplace. Psychol Sci Public Interest 2004;5(3):69–106. https://doi.org/10.1111/j.1529-1006.2004.00018.x.

[52] Moorman RH, Podsakoff PM. A meta-analytic review and empirical test of the potential confounding effects of social desirability response sets in organizational behaviour research. J Occup Organ Psychol 1992;65(2):131–49. https://doi.org/10.1111/j.2044-8325.1992.tb00490.x.

[53] Goe L, Bell CA, Little O. Approaches to evaluating teacher effectiveness: A research synthesis. National Comprehensive Center for Teacher Quality 2008.

[54] Lawless KA, Pellegrino JW. Professional development in integrating technology into teaching and learning: Knowns, unknowns, and ways to pursue better questions and answers. Rev Educ Res 2007;77(4):575–614. https://doi.org/10.3102/0034654307309921.

[55] Siddiq F, Hatlevik OE, Olsen RV, Throndsen I, Scherer R. Taking a future perspective by learning from the past—A systematic review of assessment instruments that aim to measure primary and secondary school students' ICT literacy. Educ Res Rev 2016;19:58–84. https://doi.org/10.1016/j.edurev.2016.05.002.

[56] Brantley-Dias L, Ertmer PA. Goldilocks and TPACK: Is the construct 'Just right? J Res Technol Educ 2013;46(2):103–28. https://doi.org/10.1080/15391523.2013.10782615.

[57] Schmidt-Crawford D, Tai SD, Yi Jin WW. Understanding teachers' TPACK through observation. In: Herring MC, Koehler MJ, Mishra P, editors. Handbook of technological pedagogical content knowledge (TPACK) for educators. Routledge; 2016. p. 107–18.

[58] Hofer MJ, Grandgenett N, Harris JB, Swan K. Testing a TPACK-based technology integration observation instrument. In: Koehler MJ, Mishra P, editors. Proceedings of the society for information technology & teacher education international conference 2011; 2011. p. 4352–9.

[59] Bradlow ET, Hoch SJ, Wesley Hutchinson J. An assessment of basic computer proficiency among active internet users: Test construction, calibration, antecedents and consequences. J Educ Behav Stat 2002;27(3):237–53. https://doi.org/10.3102/10769986027003237.

[60] Richter T, Naumann J, Horz H. Eine revidierte Fassung des Inventars zur Computerbildung (INCOBI-R). Z Pädagog Psychol 2010;24(1):23–37. https://doi.org/10.1024/1010-0652/a000002.

[61] Senkbeil M, Ihme JM. Diagnostik von ICT Literacy: Messen Multiple-Choice-Aufgaben und simulationsbasierte Aufgaben vergleichbare Konstrukte? Vergleich der Testergebnisse zweier Instrumente aus den aktuellen Large-Scale-Studien ICILS 2013 und NEPS. Diagnostica 2020;66(3):147–57. https://doi.org/10.1026/0012-1924/a000243.

[62] Wagener D. Der Computerwissenstest CWIS-4: Befunde zur Reliabilität und Validität. Z Personalpsychol 2003;2(4):169–81. https://doi.org/10.1026//1617-6391.2.4.169.

[63] Senkbeil M, Ihme JM, Wittwer J. Entwicklung und erste Validierung eines Tests zur Erfassung technologischer und informationsbezogener Literacy (TILT) für Jugendliche am Ende der Sekundarstufe I. Zeitschrift für Erziehungswissenschaft 2013;16(4):671–91. https://doi.org/10.1007/s11618-013-0446-5.

[64] Statista. (2022, May 5). Betriebssysteme—Marktanteile weltweit bis März (2022) | *Statista*. https://de.statista.com/statistik/daten/studie/157902/umfrage/marktanteil-der-genutzten-betriebssysteme-weltweit-seit-2009/.

[65] Petko D. Quo vadis TPACK? Scouting the road ahead. EdMedia + Innovate Learning. Association for the Advancement of Computing in Education (AACE); (2020). p. 1349–1358.

[66] de Jong T, Ferguson-Hessler MGM. Types and qualities of knowledge. Educ Psychol 1996;31(2):105–13. https://doi.org/10.1207/s15326985ep3102_2.

[67] Hurrell D. Conceptual knowledge OR procedural knowledge or conceptual knowledge AND procedural knowledge: Why the conjunction is important to teachers. Aust J Teach Educ 2021;46(2):57–71. https://doi.org/10.14221/ajte.2021v46n2.4.

[68] Baker FB, Kim SH. The basic of item response theory using R. Springer; 2017. https://doi.org/10.1007/978-3-319-54205-8.

[69] Zwitser RJ, Maris G. Ordering individuals with sum scores: The introduction of the nonparametric Rasch model. Psychometrika 2016;81(1):39–59. https://doi.org/10.1007/s11336-015-9481-x.

[70] Grossman P, Hammerness K, McDonald M. Redefining teaching, re-imagining teacher education. Teach Teach 2009;15(2):273–89. https://doi.org/10.1080/13540600902875340.

[71] Drossel K, Eickelmann B, Schaumburg H, Labusch A, Eickelmann B. Nutzung digitaler Medien und Prädiktoren aus der perspektive der Lehrerinnen und Lehrer im internationalen Vergleich. In: Bos W, Gerick J, Goldhammer F, Schaumburg H, Schwippert K, Senkbeil M, Vahrenhold J, editors. ICILS 2018 #Deutschland computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im zweiten internationalen Vergleich und Kompetenzen im Bereich computational thinking. Waxmann; 2019. p. 205–40.

[72] Whetzel DL, McDaniel MA. Situational judgment tests: An overview of current research. Hum Resour Manag Rev 2009;19(3):188–202. https://doi.org/10.1016/j.hrmr.2009.03.007.

[73] Sailer M, Stadler M, Schultz-Pernice F, Franke U, Schöffmann C, Paniotova V, et al. Technology-related teaching skills and attitudes: Validation of a scenario-based self-assessment instrument for teachers. Comput Human Behav 2021;115:106625. https://doi.org/10.1016/j.chb.2020.106625.

[74] Morgan DL. Focus groups. Annual review of sociology. (1996); 22(1):129-152. 10.1146/annurev.soc.22.1.129.

[75] Fütterer T, Steinhauser R, Zitzmann S, Scheiter K, Lachner A, Stürmer-Schröpfer K. (November 17). OSF project belonging to the publication "A test to assess teachers' knowledge about how to operate technology". https://doi.org/10.17605/OSF.IO/TGYAX.

[76] Lenzner T, Neuert C, Otto W. Kognitives Pretesting. GESIS survey guidelines. https://doi.org/10.15465/GESIS-SG_010.

[77] R. Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; 2021. http://www.R-project.org.

[78] Mair P, Hatzinger R, Maier MJ, Rusch T, Mair MP. Package "eRm". extended rasch modeling. R Foundation; 2016.

[79] Pohl S, Gräfe L, Rose N. Dealing with omitted and not-reached items in competence tests: Evaluating approaches accounting for missing responses in item response theory models. Educ Psychol Meas 2014;74(3):423–52. https://doi.org/10.1177/0013164413504926.

[80] Wright BD, Masters GN. Rating scale analysis. Mesa Press; 1982.

[81] Wright BD, Linacre JM. Reasonable mean-square fit values. Rasch Measur Trans 1994;8:370.

[82] Linacre JM. What do infit and outfit mean-square and standardized mean? Trans Rasch Meas 2002;16(1):878.

[83] Taber KS. The use of Cronbach's Alpha when developing and reporting research instruments in science education. Res Sci Educ 2018;48(6):1273–96. https://doi.org/10.1007/s11165-016-9602-2.

[84] Franke, U., Backfisch, I., Fabian, A., Breil, P., Scheiter, K., & Lachner, A. (2023, August). "I know I don't know" - Does metacognitive accuracy moderate the validity of TPK self-assessments?" EARLI 2023. Thessaloniki, Greece. [presentation].

[85] Hughes, J., Thomas, R., & Scharber, C. (2006). Assessing technology integration: The RAT – Replacement, Amplification, and Transformation—Framework. In: Crawford, Carlsen, McFerrin, Price, Weber, & Willis (Eds.), *Proceedings of SITE 2006—Society for Information Technology & Teacher Education International Conference* (pp. 1616–1620). Association for the Advancement of Computing in Education (AACE).

[86] Wirtz M, Caspar F. Beurteilerübereinstimmung und Beurteilerreliabilität. Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen. Hogrefe; 2002.

[87] Kunter M, Leutner D, Terhart E, Baumert J. Educational science knowledge and acquisition of professional competencies of prospective teachers (BilWiss). Bildungswissenschaftliches Wissen und der Erwerb Professioneller Kompetenz in der Lehramtsausbildung (BilWiss). [Data set]. IQB. Institute for Educational Quality Improvement; 2017. https://doi.org/10.5159/IQB_BILWISS_V4.

[88] Rosseel Y. Lavaan: An R package for structural equation modeling. J. Stat. Softw. 2012;48(2):1–36.

[89] Brown TA. Confirmatory factor analysis for applied research. 2nd ed. The Guilford Press; 2015.

[90] Andrich D. Rasch models for measurement. Sage Publications; 1988.

[91] Andrich D, Marais I. A course in rasch measurement theory: Measuring in the educational, social and health sciences. Springer; 2019. https://doi.org/10.1007/978-981-13-7496-8.

[92] Hauke J, Kossowski T. Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data. QUAGEO 2011;30(2):87–93. https://doi.org/10.2478/v10117-011-0021-1.

[93] Graham JW, Cumsille PE, Elek-Fisk E. Methods for handling missing data. In: Weiner IB, editor. Handbook of psychology. John Wiley & Sons, Inc; 2003. https://doi.org/10.1002/0471264385.wei0204.

[94] Bentler PM. On the fit of models to covariances and methodology to the bulletin. Psychol Bull 1992;112(3):400–4. https://doi.org/10.1037/0033-2909.112.3.400.

[95] Hu L, Bentler PM. Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. Psychol Methods 1998;3(4):424–53. https://doi.org/10.1037/1082-989X.3.4.424.

[96] Bagozzi RP. Structural equation models are modelling tools with many ambiguities: Comments acknowledging the need for caution and humility in their use. J. Consum Psychol 2010;20(2):208–14. https://doi.org/10.1016/j.jcps.2010.03.001.

[97] Cohen J. Statistical power analysis for the behavioural sciences. 2nd ed. Erlbaum; 1988.

[98] Akyuz D. Measuring technological pedagogical content knowledge (TPACK) through performance assessment. Comput Educ 2018;125:212–25. https://doi.org/10.1016/j.compedu.2018.06.012.

[99] Gerick J, Eickelmann B, Steglich E. Abschlussbericht zur prozessbegleitenden Evaluation der Einführung von Logineo NRW an Pilotschulen. Universität Paderborn; 2019. https://kw.uni-paderborn.de/institut-fuer-erziehungswissenschaft/arbeitsbereiche/schulpaedagogik/forschung/forschungsprojekte/prozessbegleitende-evaluation-der-einfuehrung-von-logineo-nrw-an-pilotschulen.

[100] Desimone LM, Garet MS. Best practices in teachers' professional development in the United States. Psychol Soc Educ 2015;7(3):252–63. https://doi.org/10.25115/psye.v7i3.515.

[101] Voss T, Kunter M, Baumert J. Assessing teacher candidates' general pedagogical/psychological knowledge: Test construction and validation. J. Educ. Psychol. 2011;103(4):952–69. https://doi.org/10.1037/a0025125.