Article

Assessing the Selection of Digital Learning Material: A Facet of Pre-Service Teachers' Digital Competence

Peter Gonscherowski * , Edith Lindenbauer , Kai Kaspar , Benjamin Rott

Article

Assessing the Selection of Digital Learning Material: A Facet of Pre-Service Teachers' Digital Competence

Peter Gonscherowski [1],*, Edith Lindenbauer [2], Kai Kaspar [3] and Benjamin Rott [1]

[1] University of Cologne, Institute for Didactics of Mathematics

[2] University College of Education Upper Austria, Department of Mathematics Education

[3] University of Cologne, Faculty of Human Sciences, Department of Psychology

* Correspondence: peter.gonscherowski@uni-koeln.de

Abstract: Given the increasing digitalization of education and the variety of available digital learning materials (dLMs) of differing quality, (pre-service) teachers must develop the ability to select appropriate dLMs for their learning goals. Therefore, educational institutions should provide pre-service teachers with opportunities to cultivate this ability. Objective, reliable, and valid assessment instruments based on established models in teacher education are necessary to evaluate the effectiveness of its development. This paper conceptualized items for assessing „selecting dLM" using a four-item straightforward yet effective and economical approach based on a model of teachers' digital competence and the TPACK framework. The scientific quality of the method was evaluated in study 1 ($n = 164$) and tested in a subsequent second, large-scale study ($n = 395$) with pre-service mathematics teachers from two universities. The empirical results indicate that the method could objectively and reliably gauge different levels of „selecting dLM" as one facet of digital competence. Furthermore, the findings obtained with the items yield small but different diagnostic results compared to a TPACK self-report instrument. The proposed approach allows for variations and integration of diverse dLMs and has the potential to be adapted in other subject areas and contexts.

Keywords: digital competence; pre-service teachers; student-teacher evaluation; technological advancement; decision-making skills; digital learning material

1. Introduction

The development of technology and technology-enhanced learning materials has transformed teaching and learning during the last decades [1,2]. This process of digitalizing education—from primary school to university level [3,4]—has been accelerated by the COVID-19 pandemic [5–8] to the point where technology has become ubiquitous in education [4,9,10]. As the term „technology" in education is ambiguous [11,12], we understand technology as software applications supporting educators' activities like organizing their work, representing learning content, and supporting self-regulated learning or collaboration. Furthermore, the same activities are supported by digital learning materials (dLMs) [13,14]. We use the term digital resources (dR) to address both digital technology and dLM [15].

Teachers play an essential role in successfully integrating dR into education. They must be competent in choosing digital technology and dLMs [1,13,16] to provide learners with appropriate learning opportunities [17]. In addition, as part of digital competence, they should know how to create digital learning content and be competent and confident in using it [12,18]. However, selecting dR is an essential skill for teachers because they have to deal with the rapid development of dR in education. Second, because of the number of available digital technologies [12,19,20] and particularly the vast number of dLMs whose quality and suitability they have to assess [21,22]. This requires objective, reliable, and valid assessment instruments to evaluate the effectiveness of developing this

facet of digital competence during the initial teacher education [16]. The knowledge required to integrate dR in the classroom is described in the TPACK framework [23], which is also often used for its assessment [15,24,25].

Two concerns arise regarding assessing the facet of selecting dR and the effort needed for its assessment. On the one hand, existing assessment instruments for teachers' digital competence, for example, require participants to evaluate the value of learning apps [21], the interpretation of video vignettes [26,27], the use of static learning material [28], or the evaluation of developed lesson plans incorporating dRs [29–31], but none directly address the selection of dR and instead focus on the overall lesson planning aspect or the mediating factors when assessing digital competence overall [15,32]. On the other hand, most instruments, especially those regarding the frequently used TPACK framework [32,33], utilize self-report assessment and, thus, lack validity and subject specificity [20,34,35] and may suffer from social desirability bias [36,cf. 37].

Because of the importance of selecting dR and addressing the concerns raised by existing self-report assessments and the lack of objective open-text items for its assessment [15], we focus this paper on developing open- and closed-text items to capture an objective approach to assess (pre-service) teachers' ability to select dR as a facet of digital competence and TPACK [15,23]. In this context, we propose assessing selecting dR using open-text items in which respondents must reason for or against selecting a given dLM for a specific learner's age, special learner's needs, and explicit learning content. We elected the context of dLM, as it has not yet been addressed methodically or empirically using open-text items in larger studies [15]. The developed items for „selecting dLM" and scoring can be used with diverse dLMs and have the potential to be adapted to other teaching subjects.

The paper's aim is threefold: first, to elucidate the genesis of items for objectively assessing the digital competence facet of „selecting dLM" for pre-service teachers and its position in existing theoretical competence models; second, to empirically validate the developed items (study 1); and third, in a larger follow-up study (study 2), to compare the results of the developed items for objectively assessing „selecting dLM" with those of an established TPACK self-report instrument by Schmid et al. [38] that also addresses this facet of digital competence.

2. Theoretical Background

2.1. Teacher Competence in Teacher Education

The professional competence of pre-service teachers is an outcome of teacher education that is dispositional and situation-specific and consists of cognitive and affective-motivational factors, including situation-specific skills [39–41]. The level of professional competence acquired during teacher training is primarily determined by the didactic quality of teacher training, the opportunities to learn [41], and prospective teachers' learning. Teacher education enables pre-service teachers to master professional tasks like lesson planning [16]. As part of lesson planning, selecting appropriate dR is one of the tasks contributing to student learning outcomes [17,42]. The availability and affordances of dRs in education [18] have changed all aspects of teacher education--teacher education itself, the professional competence of teachers, the process of developing said professional competence, and the consequential learning outcome on learners—in the teacher education effectiveness model by Kaiser and König [16,43].

The TPACK model [23,44] describes what pre-service teachers need to know to integrate dRs in teaching successfully.

2.2. TPACK Framework

Figure 1 shows the TPACK framework [23], which expands the seminal PCK framework of Shulman [45] by adding technological knowledge and the resulting intersections. It describes the knowledge teachers require to successfully integrate technology in teaching as technological, content, and pedagogical knowledge and the respective interactions of these knowledge components. Namely

technology content knowledge (TCK), technology pedagogical knowledge (TPK), pedagogical content knowledge, and technology pedagogical content knowledge (TPACK).

Technological knowledge is conceptualized as developmental knowledge evolving according to technological changes. It encompasses teachers' knowledge, enabling them to accomplish various tasks using technology. TCK refers to the reciprocal relationship between content and technology. Teachers must learn how technology can alter—enhance or hinder—the learning of the content and how integrating technology can change the subject matter. TPK refers to technology integration in teaching and learning processes and the knowledge of how various technologies impact such processes. TPACK, the key component that brings everything together, encompasses teachers' knowledge of challenges and changes in teaching when using technology, factors that make content-related concepts easier or more challenging to learn, and ways to overcome learning difficulties using technology. TPACK is the basis for sound, meaningful teaching with technology [23]. Content knowledge is described as knowledge of the learning content, the curriculum knowledge regarding the learning content, and the special needs of learners [45]. Literature suggests that content knowledge is essential when selecting dRs [15,18,43].
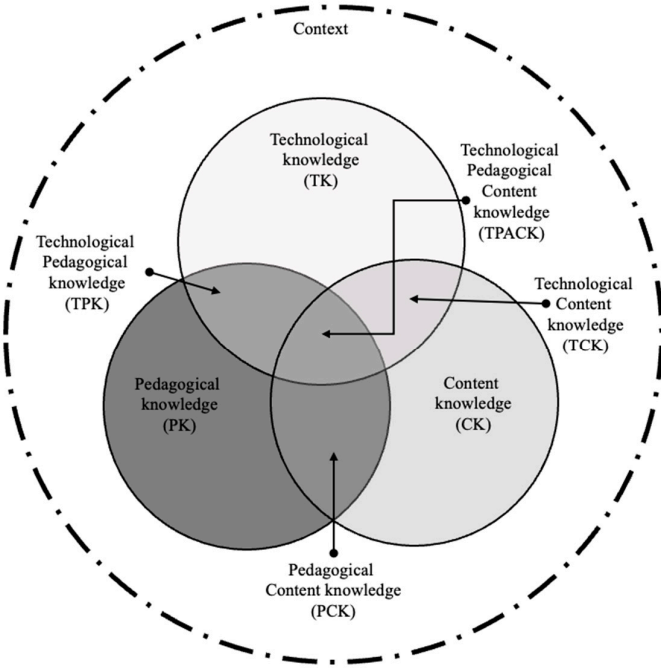

Figure 1. TPACK framework adapted from Koehler et al. [44].

Although the TPACK model's naming and terms are described as knowledge, the research community understands that they are more than knowledge and entail attitude and skill, which, therefore, describe competence [32,39,43,46,47]. Nevertheless, we follow the naming convention established by Koehler et al. [24] and refer to the framework and the intersection of all three components, content-, pedagogical-, and technological- knowledge, as TPACK.

The TCK, TPK, and TPACK components and the external context-specific factors describe the knowledge teachers require to integrate technology successfully in their teaching. The required knowledge described in the TPACK framework can also be used to justify and reason for selecting dRs. In addition, the reasoning used in the dRs selection process can be structured using the TPACK components, namely subject-specific TCK reasons, subject-unspecific pedagogical TPK reasons (see sections 2.2.1 and 2.2.2), and TPACK reasons. In the following, TCK-x and TPK-x numerically denote the various TCK and TPK reasons, as listed in Table 1. The TCK-x and TPK-x reasons were identified in the qualitative interview study we conducted with pre- and in-service teachers' reasoning when

selecting dRs [48]. Section 3 applies then the reasoning for selecting dR specifically to „selecting dLM."

2.2.1. TCK-Related Reasons for Selecting Digital Resources

TCK refers to the reciprocal relationship between content- and technological- knowledge and thus is teaching subject-specific [23]. Because of the importance of mathematics in general and its many connections to other subjects, especially STEM subjects, this section focuses on mathematics-relevant reasoning.

In mathematics education, dRs provide new, dynamic, and different representations [19,49–52]; thus, those would be reasons for selecting dRs. Using dRs, multiple representations can be dynamically linked in new ways [53]. Another reason for integrating dRs into mathematics education is their potential to reduce or modify extraneous cognitive load [48,54] (Table 1, TPCK-5/6) and outsource repetitive tasks, saving lesson time and allowing students to focus more on higher levels of mathematical thinking, leading to better learning outcomes [12,18,50,54]. However, dRs can also increase the extraneous cognitive load if learners are unfamiliar with them or dRs are inadequate for the learning content (Table 1, TCK-4). TCK reasoning can be summarized as enabling different (real-world) representations (Table 1, TCK-1/2), dynamic representations (Table 1, TCK-3), and decreasing, increasing, or modifying extraneous cognitive load (Table 1, TCK-4/5/6), with each reasoning abbreviated by TCK-x (Table 1).

Other teaching subjects can entail additional TCK reasoning specific to them but not applicable to mathematics. Examples would be the ability to simulate phenomena invisible to the eye in biology or chemistry [27,55], connecting physical phenomena with their representations in physics [52], or understanding content through audio cues in language education [56]. Section 3.2 describes the TCK-x reasons in the context of a specific dLM, thus TCK-x reasoning for „selecting dLM."

2.2.2. TPK-Related Reasons for Selecting Digital Resources

TPK refers to technology integration in teaching and learning processes and is subject-unspecific, as is TPK reasoning when selecting dRs. We denote each TPK reasoning numerically as TPK-x, as referenced in Table 1. dRs can potentially support students' self-regulated learning processes [57–60]. However, this requires teachers to have the appropriate skills to orchestrate the use of technology in the classroom to ensure learners are using technology to achieve the intended learning aim (Table 1, TPK-6) and are not distracted by it [57]. Next, dRs enable teachers to design lessons that enhance discovery learning (Table 1, TPK-3), keeping students engaged while exploring new topics [21,61]. The importance of learner engagement and motivation may also be factors when teachers choose a dR (Table 1, TPK-1) [56,57,59,62]. The latter is linked to student learning outcomes [12,18,42,50,63]. Applying dRs in assessments may positively impact teachers' efficiency (Table 1, TPK-5) and can be a factor in the decision to use them [33,48]. From a learner's perspective, it is essential to consider the accessibility of dLMs to the learners and their individual needs and whether the technology can promote differentiation and inclusion (Table 1, TPK-4) [19,61,63]. To recapitulate, TPK reasoning when selecting a dR can be learner motivation (TPK-1), self-regulated learning (TPK-2), discovery learning (TPK-3), differentiation and inclusion (TPK-4), increasing teacher efficiency (TPK-5) and distracting to learners (TPK-6). Section 3.2 describes the TPK-x reasons in the context of a specific dLM, thus TCK-x reasoning for „selecting dLM."

Table 1. Coding model for TCK-x, TPK-x, and TPACK reasoning.

TPACK Comp.	Description of code (#)	DigCompEdu [63], TPACK [23], teaching core practices Grossman [61]	Studies supporting the code
TCK-x	Different/real world representations (TCK-1/2)	„newer technologies often afford newer and more varied representations" [23] (p. 1028)	[19,52]

	Dynamic representation (TCK-3)	-	[18,64]
	In-/decrease, Modify extraneous cognitive load (TCK-4/5/6)	-	[48,54]
TPK-x	Motivation (TPK-1)	„... Teachers design and sequence lessons [...] while keeping students engaged...” [61] (p. 167)	[56,59,60,65]
	Self-directed learning (TPK-2)	„To use digital technologies to support learners' self-regulated learning” [63] (p. 21)	[18,60,66]
	Try out, explore, discover (TPK-3)	„...digital technologies can be used to facilitate learners” [...], e.g., when exploring a topic…” [63] (p. 22)	[18,21,61]
	Differentiation, inclusion (TPK-4)	„...To ensure accessibility to learning resources and activities, for all learners, including those with special needs…” [63] (p. 22)	[4,19,21,42]
	Teacher efficiency (TPK-5)	„...composing and selecting assessments, teachers consider validity, fairness, and efficiency.” [61] (p.168)	[33,48,60]
	Distracting to learners (TPK-6)	-	[48,57]
TPACK	Combination of TCK, TPK	„TPCK is the basis of good teaching with technology [...] pedagogical techniques that use technologies in constructive ways…” [23] (p. 1029)	[48,67]

In summary, the reasoning for or against a dR in a particular teaching situation can be multifold and depend on the situation (learner age group, special educational needs) and context (infrastructure).

It can be justified using TCK, TPK, or a combination of both, referred to as TPACK reasoning. Table 1 comprehensively lists the TCK-x, TPK-x, and TPACK reasoning. There is not necessarily a single decision to select a digital resource for different teaching situations and contexts and a specific digital resource implementation of a particular learning content. Thus, the reasoning behind its selection, not the decision itself, is essential for assessing „selecting dR.”

Considering the multifaceted reasoning involved in selecting dRs and the significance of this facet of digital competence for the development of pre-service teachers, we also recognize the methodological and empirical research gap concerning objective, reliable, and valid open-text items for its assessments [15]. This gap is particularly evident in the context of dLM and „selecting dLM,” which leads us to pose the following research questions.”

RQ1: How can „selecting dLM” be assessed using open-text items?

In addition to the validation study (study 1) for the items developed in RQ1 to assess „selecting dLM,” we apply them in an online test with a larger sample of pre-service mathematics teachers (study 2). To gain a deeper understanding of the items we have developed, we compare results from objectively evaluating the performance of „selecting dLM” using the items we created with those obtained from a TPACK self-report instrument developed by Schmid et al. [38]. Based on this, we pose the following research questions.

RQ2: What insights can be gained using the developed items with a larger sample of pre-service mathematics teachers?

RQ2.1: Can the developed items for „selecting dLM" assess different levels regarding the number of semesters of study, and are the results distinguishable from TPACK self-report results?

RQ2.2: What is the relationship between the developed items?

RQ2.3: What specific reasoning is considered when selecting a dLM?

This section outlines the development of the items for assessing „selecting dLM." It includes a detailed description of one dLM (dLM1) and the reasoning used in its selection. In addition to a summary of other dLMs (dLM2-dLM4) and the applied scoring methods. Furthermore, it presents the two studies and the participant samples used to address research questions 1 (RQ1, study 1) and 2.x (RQ2.x, study 2).

3. Materials and Methods

3.1. Genesis of Items for Assessing the Skill of „Selecting dLM"

The items were constructed using a multi-methods approach. At the process level, we conducted a systematic literature review of existing TPACK assessment methods for assessing the facet of digital competence of „selecting dR" [15,32]. In addition, we engaged in individual and group discussions with mathematics teacher educators and in-service teachers to elucidate what digital resources they use in their teaching and their reasons for selecting them.

Furthermore, we used small pilot studies testing different open-text items and contexts [67–69], supplemented by a qualitative interview study with pre- and in-service teachers, reasoning for or against using digital resources for a specific learning content and learner group [48].

At the conceptual level, the developed items shown in Table 2 are based on the competence model in teacher education by Blömkete et al. [39], distinguishing between the reasoning for „selecting dLMs" as a cognitive disposition, the essential content knowledge needed for decision-making [18], and the situational factors influencing reasoning in such decisions [39]. The specifications for essential knowledge are derived from the content knowledge definition within the TPACK framework [23,45], as outlined in items 1-3 in Table 2. The perception and interpretation of situational and contextual factors in decision-making are identified as TCK, TPK, and TPACK reasoning, found in item 4 (see Table 2). The resulting four items present a very economical approach, assessing „selecting dLM" with only two open and two closed items.

Table 2. Items used to assess „selecting dLM.".

Item no.	TPACK component	Item wording	Type of item
1		For which learner age do you think the presented digital material is suitable?	single-choice selection of a grade level[1]
2	Content knowledge	In your opinion, is the presented digital material suitable for learners with special educational needs? If so, select one or none.	single-choice selection of (no) special need[2]
3		Describe the learning content for which you think the presented digital material is intended.	open-text item
4	TCK-x, TPK-x, and/or TPACK reasoning	For the specified grade level, special educational needs, and your description of the learning content of the presented dLM, justify why or why not you would use the presented digital material.	open-text item

[1] To capture the learner's age, we provided a list of grade levels in increments of two grade levels aligned with the local curriculum. The following options were given:1-2, 3-4, 5-6, 7-8, 9-10, 11-13. [2] To capture the special needs of learners, we provided a list of special needs designations aligned to the local terminology in Austria and Germany. The following options were given: not suitable for learners with special needs, social & emotional

needs, mental development needs, hearing and communication needs, motoric needs, learning needs, speech needs.

3.2. Specific dLMs Used in the Studies for Evaluating the Developed Items

To provide context for evaluating the items for „selecting dLM," we have chosen multiple dLMs within GeoGebra because of its wide distribution, that is, 100 million users in 190 countries, and open-source access to more than one million open, accessible dLMs [70]. For dLM1 (Figure 2), we explain in detail our decision-making process and why we used it to evaluate the items. The same process was used for dLM 2, dLM3, and dLM4, also used for evaluating the items.
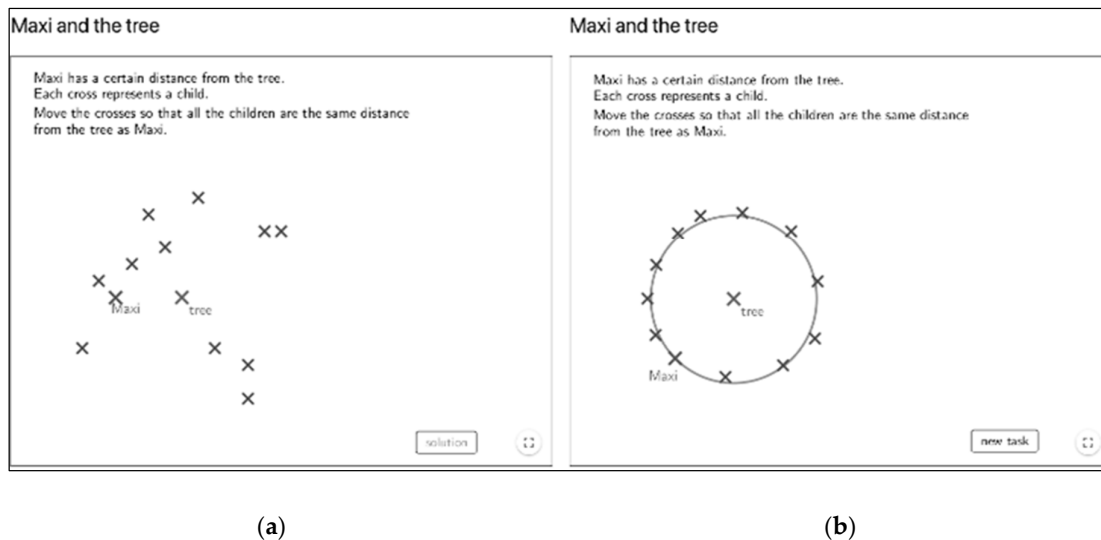

(**a**)   (**b**)

Figure 2. (**a**) Shows the starting point of the dLM, and (**b**) shows the dLM after pressing „solution." For this publication, the dLM was translated from German and adapted for printing by the authors.

The dLM 1 (Figure 2) explores the mathematical concept of a circle. It is suitable for primary and secondary school students and accessible to students with special educational needs. Iconic and enactive non-digital alternative materials exist to explore the learning content. Thus, the dLM allows for diverse reasoning for or against its selection. As illustrated in Figure 2a, using the dLM1, learners should discover a circle as a concept based on its defining property: a shape composed of all the points in the plane that are a certain distance (radius) from a certain point (the center). The learning material consists of an introductory text, a task, and various crosses representing a tree (dark-colored cross), a child named Maxi (dark-colored cross), and several other children (light-colored crosses). Figure 2a shows Maxi at a pre-set distance from the tree. Now, the learners should move the light-colored crosses to be the same distance from the tree as Maxi and discover that they form a circle. A circle appears once the learners push the „solution" button, as shown in Figure 2b, reinforcing the concept of the circle and its defining property of all points (Maxi and other children) on the circle having the same distance from the center (tree).

Possible arguments for or against this dLM are that learners can self-regulate (Table 1, TPK-2), discover (Table 1, TPK-3), and check whether their solution is correct. The shortcomings of the dLM are that learners can place all crosses (children) on the same point and that not all crosses (children) need to be moved before pressing the solution button. The latter corresponds to increasing extraneous cognitive load, which can negatively impact learning (Table 1, TCK-5). The knowledge of these shortcomings is a reason for not selecting the dLM.

However, in contrast, one can argue that the dLM reduces the extraneous cognitive load [50,54,66], as the abstract concept of a circle and its defining property is made available in a virtual, enactive way (Table 1, TCK-1/2). In addition, the cognitive activity of manually constructing and

drawing a circle is outsourced to the dLM, leading to reduced extraneous cognitive load (Table 1, TCK-5.) Learners manipulate the crosses using their fingers or mouse. Thus, using the dLM gives learners dynamic control of the presented objects, equivalent to a TCK affordance (TCK-3).

The outlined affordances of dLM 1, the learning content it is intended to deliver, and its place according to the curriculum are specific to this particular dLM. They are in the context of the local curriculum. In addition to the described dLM1, which was used in studies 1 and 2, dLM2: symmetry and axis of symmetry [71], dLM3: generate bar charts [72], and dLM4: arithmetic mean [73] were used in study 1 only. The scoring used for all dLMs for items 1-3 and item 4 is outlined next, using dLM1 again as an example.

3.3. Scoring of the Items Used for Assessing the Skill of „Selecting dLM"

The understanding of the learning content, the suitability for a learner's age, and the special needs of learners (Figure 3, items 1-3) were evaluated together, as varying reasoning is possible depending on the former. We applied a four-point scale ranging from zero to three. The authors derived the scoring scale partly inductively from the responses and deductively from the German and Austrian curricula [74,75]. The latter outlines the requirements for learners with special needs and the appropriateness of learning content for a specific learner's age.

When no description or an incorrect definition of the learning content was provided, the responses received a score of zero. A score of one was awarded for describing the learning content correctly but inappropriate for the learner's age group or the learner's particular educational needs specified in items 2 and 3 (Figure 3). A generic description of the learning content that did not mention the property of the circle as the equal distance from the center to every point on the circle received a score of two, while a score of three was awarded if the property was mentioned. An example response for the former states, „Elaboration of the topic circle (diameter, radius) for learners aged 7-8 with specific motor, emotional, and social needs." An example for the latter states, „To introduce the circle and draw the learner's attention to the fact that every point on the circle is exactly equidistant from the center point" for learners aged 5-6 with no special educational needs. The same approach was applied to score the responses to items 1-3 for the other dLMs respective to their learning content. Table A1 in the appendix provides examples of the scoring of all four dLMs (dLM1- dLM4).
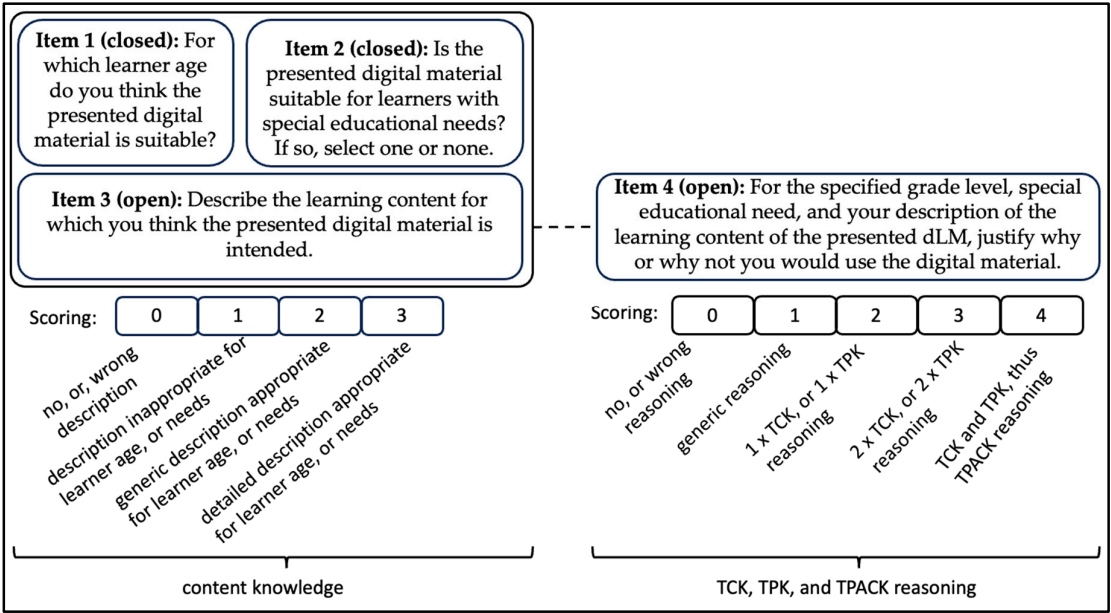

Figure 3. A graphical representation of the structure and the scoring of items 1-4.

The reasoning for selecting the dLM (Figure 3, item 4) was scored from zero to four on a 5-point scale. Accordingly, the reasonings were applied as previously categorized as TCK-x and TPK-x reasoning (Table 1). The scoring scale was inductively derived from the responses. A score of zero was given when no reasoning was provided and a score of one for generic reasoning. If a response included one TCK-x or one TPK-x reason in the justification, a score of two was awarded, and when two TCK-x or two TPK-x reasons were included, a score of three was given. If the participant's reasoning included both TCK-x and TPK-x reasoning, we considered it TPACK reasoning and scored it with four. An example of TPACK reasoning is as follows: „...[the dLM] enables enactive and visual learning that actively involves learners in the learning process. Through the concrete task of positioning the cross in a circle around the tree, the children experience geometric concepts such as the radius, the center, and the circular shape...The combination of movement, practical experience, and reflection makes learning sustainable and motivating, appealing to different learners..." This response was provided in the context of the learner's age, grade levels 3 and 4 (Figure 3, item 1), and no special needs of learners (Figure 3, item 2).

It is worth noting that, except for TPK-5 reasoning, justifying the selection to increase the teacher efficacy, the learner's age and special needs were considered in the reasoning scoring (Figure 3, item 4). Reasoning, such as increasing or decreasing the extraneous cognitive load using a dLM, depends on the specified learner's age and special needs for learning content. The dashed line in Figure 3 between items 1-3 and 4 presents this relation. Table A2 in the appendix provides scoring examples for all four dLMs (dLM1- dLM4) for item 4.

3.4. Participants of Studies 1 and 2

We conducted two separate studies to address the research questions; see Table 3. In both studies (study 1 and study 2), we employed convenience sampling, which included pre-service mathematics teachers from universities in Austria and Germany. We selected universities in these countries because they share a common language and have similar curricula. Participants were informed about the purpose of the study, and their participation was voluntary.

In the first study (study 1), we focused on establishing the scientific quality of the items we have developed to assess „selecting dLM" (RQ1). The study included four dLMs (dLM1-dLM4), as listed in Table 3 (including the references for the dLMs.) The designed items (1-4) were presented repeatedly to participants four times, each time in context with a different dLM (dLM1-dLM4).

Table 3. Initial and adjusted sample sizes for both studies.

Study	RQs	Size of initial sample	Size of adjusted Sample[1]	Adjusted sample size per university of country: Austria / Germany	dLM used in study	Mean processing time of task in minutes
1	RQ1	164	164	61 / 103	dLM1-4	9.84[2]
2	RQ2.x	395	379	55 / 324	dLM1	4.29[3]

[1] In the first study (study 1), we included all responses to evaluate the robustness of the coding and scoring approach and establish a lower time limit for carefully processing the items (2.50 min). In study 2, all responses that did not meet the established minimum processing time were removed. [2] dLM1-dLM4 were included: (dLM1) circle and the property of the radius [76], (dLM2) symmetry and axis of symmetry [71], (dLM3) create bar charts [72] and (dLM4) arithmetic mean [73]. [3] only dLM1 was included: circle, and the property of the radius [76].

In the second study (study 2), we recruited new participants using the same enrollment approach at the same universities as in study 1. We applied the items previously validated in study 1 and self-reported TPACK through the Likert scale instrument developed by Schmid et al. [38]. The latter scale includes statements that address „selecting dLM" and stating, „I can select technologies to use in my classroom that enhance what I teach, how I teach, and what students learn." [38] (p. 4)

For test-time economics, we included only one dLM in this larger study (n = 395). The dLM (dLM1) for exploring the properties of a circle was utilized in the study, as it aligns with the curriculum and is well-suited for a broad participant pool from both universities. Additionally, iconic and enactive alternatives to the dLM exist, allowing for diverse reasoning.

The mean lower processing time for an individual dLM in study 1 (dLM1–dLM4) compared to study 2 (dLM1 only) may be attributed to the repetitiveness of the task and familiarity with the items when evaluating four dLMs (dLM1-dLM4).

4. Results

To answer the research questions, we used R, version 4.4.2 (2024-10-31), for statistical analysis and the package psych, version 2.4.12. All statistical testing used a signification level of 5%, and the appropriate effect size statistics were established [77].

4.1. RQ1: How Can „Selecting dLM" Be Assessed Using Open-Text Items?

To evaluate the developed items' scientific quality, we used the criteria for psychometric instruments: objectivity, reliability, and validity [78] for the sample of mathematics pre-service teachers in study 1 (n = 164) and dLM1-dLM4.

To evaluate the objectivity of the coding and scoring of the items, we employed two coders who independently categorized the participants' responses. The open-text responses to item 3 were evaluated and scored on a scale from zero to three based on the selections made in items 1 and 2 (see Figure 3). The open-text item 4, representing TCK, TPK, and TPACK reasoning, was coded according to the coding model in Table 1 and scored on a scale from zero to four (see Figure 3). Differences in the coding and scoring were discussed and resolved. During these discussions, examples of the coding, scoring (Tables A1 and A2), and coding model (Figure 3) were used as guidelines.

To assess the reliability of the items, we computed Cronbach's Alpha coefficients, as items 1-4 were used four times, each time in the context of a different dLM (dLM1-dLM4.) This denoted a repetition of the measurement „selecting dLM." Using the scores established for items 1-3 across the four dLMs (dLM1-dLM4) yielded a value of .83, and for item 4, a value of .91 for Cronbach's Alpha. These coefficients indicate good to excellent reliability, supporting the reliability of the items, the categories (Table 1), and the scoring (Figure 3) for evaluating „selecting dLM."

Content validity is provided since the designed items are based on scientifically tested and widely accepted frameworks and models, namely the TPACK framework by Mishra and Koehler [23] and the knowledge, skill, and performance model by Blömeke et al. [39]. In addition, Grossman's [61] core teaching practices were utilized to develop the reasoning categories (Table 1), reflecting a more practice-oriented and Anglo-American perspective. In contrast, DigCompEdu [63] was also employed to establish reasoning categories (Table 1) that describe a more conceptual, theoretical, and European framework for digital competence, further strengthening the approach's validity.

Moreover, content validity was enhanced by engaging peers, teacher educators, and in-service teachers in developing the items. The pilot [67,69] and interview studies [48] , which examined the reasoning behind selecting dRs in preparation for the items' genesis, further support the items' content validity.

4.2. RQ2: What Insights Can Be Gained Using the Developed Items with a Larger Sample of Pre-Service Mathematics Teachers?

To further examine the items (1-4) and the insights they provide regarding their generalizability, we applied them in an online test with a larger sample of pre-service mathematics teachers in study 2 (n = 379), using only dLM1 for reasons of test time economy. Additionally, we compared the external performance of „selecting dLM" with the developed items against the self-reported TPACK performance of the participants using the TPACK instrument by Schmid et al. [38].

For more detailed analyses, we grouped participants in increments of two semesters, as the number of semesters in the development stages varies between the two universities (Table 4). Due to the smaller subgroups of pre-service teachers by study stages (primary, special education, lower and upper secondary education), we examined differences based only on semesters of study rather than on stages of study as was done in other studies [79,80]. Following the coding and scoring we established and validated using the dataset from study 1, we employed two trained coders who independently categorized the participants' responses. The inter-coder agreement was calculated separately for coding items 1-3 (Figure 3), κ = .86, and for item four (Figure 3), κ = .99, placing them in a near-perfect range [81].

4.2.1. RQ 2.1: Can the Developed Items for „Selecting dLM" Assess Different Levels Regarding the Number of Semesters of Study, and Are the Results Distinguishable from TPACK Self-Report Results?

Using the dataset from study 2 (*n* = 379), we examined the ability of the items to differentiate different levels of „selecting dLM." We examined the scores concerning the number of semesters—using one-factor ANOVAs. Normality requirements were tested by inspecting Q-Q plots, skewness, and kurtosis [82] and using Levene's test for homogeneity of variance and the normal distribution of residuals. Because of the smaller subgroups in the sample, we did not use the universities of the two countries, Austria and Germany, or the study stages as a factor in the one-factor ANOVAs or conducted two-factor ANOVAs.

Table 4 reports the descriptive results of the pre-service teachers' scores for „selecting dLM" by number of semesters of study (external assessment) and the self-reported TPACK using the TPACK instrument by Schmid et al. [38]. Pre-service teachers achieved the highest results for „selecting dLM" in semester 7 and higher (*M* = 2.65, *SD* = 1.51).

Table 4. Descriptive statistics and ANOVA results for the external assessment of *„selecting dLM"* by the number of semesters of study and Kruskal-Wallis test results for self-report TPACK.

Type of assessment	Sem. 1, 2 (*n* = 57)		Sem. 3, 4 (*n* = 149)		Sem. 5, 6 (*n* = 71)		Sem. ≥ 7 (*n* = 102)			
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*		
external [1]	1.37[a]	1.43	2.12[b]	1.60	2.27[b]	1.55	2.65[b]	1.51	**F(3,375)** 8.51	ηp^2 .06
self-report [2]	3.12[c]	.86	3.27[c]	.73	3.44	.72	3.55[d]	.76	**Chi² (3)** 11.99	

[1] items 1-4 for assessing „selecting dLM" proposed in this paper, scale 0-8. [2] Schmid et al. Likert scale [38] (1 = *strongly disagree* to 5 = *strongly agree*). [ab] Means without a common superscript differ (*p* < .001). [cd] Means without a common superscript differ (*p* < .007).

The one-factorial ANOVAs revealed statistically significant differences in the number of semesters (see Table 4). Based on these results, we can conclude that the items differentiate different levels of „selecting dLM."

To analyze whether the developed items for „selecting dLM" for the participants in study 2 are distinguishable from the results of the TPACK self-report instrument by Schmid et al. [38], we examined the latter as well concerning the number of semesters of study using Kruskal-Wallis tests (Table 4). First, we calculated Cronbach's alpha for the TPACK self-report scale, which was .97. This indicated excellent reliability and surpassed the .87 reported in the original publication [38]. The descriptive results showed again that participants in semesters 7 and higher exhibited the highest self-reported TPACK (*M* = 3.55, *SD* = .76).

We calculated effect sizes for each assessment to compare the results of self-reported TPACK [38] and the external performance of „selecting dLM." The effect size for the „selecting dLM"

performance assessed using the items we developed was medium (f = .26). In contrast, the self-reported TPACK exhibited a small effect size (f = .23) [77].

4.2.2. RQ2.2: What is the Relationship Between Learning Content Knowledge (Items 1-3) and the Rationale (item 4) for „Selecting a dLM"?

We use Spearman rank correlation to determine whether the scores for content knowledge (items 1-3) and TCK-x, TPK-x, and TPACK reasoning (item 4) support the literature's notion that higher content knowledge parallels better reasoning [15,18].

The results reveal a positive but small correlation between pre-service teachers' content knowledge (items 1-3) and their TCK-x, TPK-x, and TPACK reasoning (item 4) (r_s = .17, p = .001).

4.2.3. What Specific Reasoning Is Considered When Selecting a dLM?

Figure 4 illustrates the frequency of the reasoning (item 4) used to justify „selecting dLM," employing the abbreviations TCK-x and TPK-x to denote specific reasoning (Table 1). The surface areas represent the frequency of each reasoning type. No justification (33.0%) and general justifications (18.7%) together account for slightly more than half of the total justifications (51.7%).

TCK reasoning (24.3%), including individual TCK-[1–6] reasoning, was used more frequently than TPK reasoning (18.7%), which includes individual TPK-[1–6] reasoning. Responses with either two TCK or two TPK, along with TPACK reasoning that integrates both TCK and TPK reasoning, were the least frequent in the sample (6.3%).

Figure 4. Frequency of no-, generic, single/multiple TCK, TPK, and TPACK reasons used to justify the selection of dLM1.

Comparing the reasoning used for dLM1 in study 1 (n = 164) and study 2 (n = 379) descriptively reveals that the same top six reason codes are utilized, though in a different order. In both studies, no and generic reasoning proportions exceeded fifty percent (study 1: 61.0% and study 2: 50.7%). This is followed by TCK-1 (different representations), TCK-5 (decreasing extraneous cognitive load), TPK-2 (self-directed learning), TCK-4 (differentiation, inclusion), and TPK-1 (motivation) reasoning, representing 28.7% in study 1 and 37.7 % in study 2 in total—other reasoning attributed for the

remainder. Study 1 and 2 participants differed in structure based on the number of semesters and the study stage.

5. Discussion

Digital competence is essential to teachers' professional competence. As part of digital competence, we identified the facet of „selecting dLM" as a crucial aspect that has yet to be addressed empirically in larger studies [15] and methodologically using open-text items. Therefore, we conceptualized the facet using the general model of teachers' competence by Blömeke et al. [39] and the TPACK framework [23] and developed and evaluated open and closed items suited to assessing it objectively.

5.1. Discussion RQ1: How Can „Selecting dLM" Be Assessed Using Open-Text Items?

The proposed approach is characterized by a suitable process since the items' operationalization and coding procedures are documented intersubjectively and comprehensibly. In addition, standardized online tests were used to capture the data consistently.

Furthermore, high inter-rater agreement in study 2 indicated the objectivity of the coding and scoring. The reliability of the items, as measured by Cronbach's alpha, was good (study 1). A critical question is whether using Cronbach's alpha to assess the reliability of the items is justified, given that the same items were used, albeit in different contexts (dLM1-dLM4). Other approaches, such as developing different items, retesting at different times, or using the existing items in more different contexts, as demonstrated in one of the pilot studies we conducted [67,69], are alternatives. However, the latter approach carries other risks, as the reasoning for selecting dLM may differ from the reasoning for selecting digital technology or the knowledge that test participants acquire over time. The variation of the dLMs used in this study presents sufficiently different contexts using the same items to assess „selecting dLM," however, the high Cronbach's alpha values may be explained by using the same items [83]. Participants in studies 1 and 2 used similar reasoning for dLM1 (section 4.2.3, research question 2.3), which is further evidence of reliability despite the participant pool being different in structure in both studies.

Moreover, using peers, teacher educators, and in-service teachers in the genesis of the items and the smaller quantitative pilot [67,69] and a qualitative interview study [48] further substantiated the approach and its validity. Content validity is further strengthened through the items' origins and using established theoretical frameworks, namely the TPACK framework by Mishra and Koehler [23] and the competence model by Blömeke et al. [39] , in the development. In terms of the competence model by Blömeke et al. [39], the items (1-3) for one assess dispositions. In particular, the dispositional knowledge of the learning content a dLM intends to deliver [15,18]. Furthermore, the items capture knowledge of where the learning content is located within the curriculum regarding the learner's age and special needs [45].

Additionally, item 4 assesses the interpretation and perception of the dLM regarding its suitability for supporting (TCK-5) or hindering (TCK-4) learners of specific ages with special needs and the identified learning content. The same applies to the other reasoning categorized in Table 1, such as describing the use of a dLM as motivational (TPK-1) or justifying the selection of a dLM based on the dynamic representation of the learning content (TCK-3). Regardless of the underlying reasoning, the items require individuals to reason based on their interpretation and perception using the reasoning identified in Table 1. Thus, they entail elements of skill as described in the competence model by Blömeke et al. [39]. Therefore, we argue that items 1-4 assess dispositional knowledge and skill. As Tabach and Trgalová noted [47], we do not claim that the items represent competence. Still, we assert that what they assess lies on the spectrum between dispositional knowledge and competence.

In addition, the literature and conceptualization support the notion that items 1-4 for assessing „selecting dLM" objectively capture TCK, TPK, and thus TPACK reasoning. In the scoring, we have established we differentiate single TCK-x or TPK-x, multiple TCK-x, and TPK-x, and TCK and TPK,

hence TPACK reasoning, but does that mean we capture TPACK? For obvious reasons, reasoning using just one TCK-x and TPK-x reason, hence TPACK reasoning (Figure 3), does not represent TPACK in its entirety. Even if we would append the scoring we have established in Figure 3 to require all TCK-x and TPK-x reasoning for obtaining the maximum score, one could still argue it is only in the context of a single dLM and thus limited. However, similar limitations apply to other assessments, such as the evaluation of lesson plans or observations, which are also limited in scope. The study by Jansen et al. [84] cites three to ten selection opportunities for a lesson plan and four to eighteen justifications. The latter equates to an average of two justifications per selection, the same number of reasoning considered in our scoring as the maximum (Figure 3). Given the economics of pre-service teacher assessments, a decision has to be made regarding the assessment type and the time required. We suggest that our approach using four items represents a very time-economical approach and thus may not be as encompassing as other more time-intensive approaches, e.g., evaluation of lesson plans.

In addition, it should be noted that „selecting dLM" is only one facet of digital competence pre-service teachers require. Other facets, such as decision-making in a (digital) classroom setting [85,86], the creation of digital learning materials [18], the variation of problem tasks [87], problem-solving [88], and information literacy [89], are also essential and require focus in teacher education. However, considering the changes involving AI-based technologies, which can potentially transform how and what is taught [90], „selecting dLM" and „selecting dR" is one of the crucial facets of the digital competence of pre-service teachers [1,13,16].

5.2. Discussion RQ2.x: What Insights Can Be Gained Using the Developed Items with a Larger Sample of Pre-Service Mathematics Teachers?

The results show that the items are sensitive to the semesters of study of pre-service teachers. Pre-service teachers with fewer semesters of study score lowest and have less ability to describe the learning content a dLM intends to deliver, place the learning content in the curriculum, and justify their selection of dLM—they have a lesser level of „selecting dLM." These results are consistent with the literature [48,80,91].

The statistical dependence of content knowledge (items 1-3) and sound reasoning (item 4) revealed by the Spearman correlation is consistent with the conceptualization of the items for assessing „selecting dLM" in the competence by Blömeke et al. [39]. The empirical results prove that the items align with the widely accepted notion that skill depends on (content) knowledge [15,18,41]. However, given the small effect size, that relation may not be as strong as assumed in the literature concerning „selecting dLM" and the four items.

The small and medium effect sizes of the ANOVAs and the Spearman correlation may be due to other factors influencing the facet of „selecting dLM." These factors include the university curriculum and coursework [21,92], previous experience [48], motivation, and attitude toward technology [46,93]. Other studies assessing digital competence using the TPACK framework, e.g., evaluating lesson plans, which entails the facet of „selecting dLM, „ also do not consistently find or report different levels of TPACK [94–96].

Comparing the results of the facet „selecting dLM," objectively assessing (one facet of) TPACK and the results of self-reported TPACK [38] show different effect sizes regarding the number of semesters. However, the differences are minor (f = .26 versus f = .23), which indicates that the external performance of „selecting dLM" parallels the self-reported TPACK performance in terms of number of semesters. This aligns with the aims of educational institutions to make pre-service teachers competent and confident in „selecting dLM" [12,18].

The objective results for pre-service teachers at both universities indicate room for improvement in „selecting dLM." Even the median scores of the subpopulations with the highest performance are less than half of the maximum. The low results for pre-service teachers at the beginning of their teacher education can be attributed to their entry dispositions [43]. Nevertheless, the overall low results for pre-service teachers in higher semesters could stem from the learning opportunities and

outcomes of teacher education regarding „selecting dLM." Furthermore, the latter could also be explained by pre-service teachers' beliefs about teaching with technology [93]. Additionally, voluntary participation in the online tests could account for some of the low results, as supported by the high percentage of „no-reasoning" in both studies (Figure 4).

5.3. Consequences for Teacher Education and the Development of Selecting dLM

We also believe that evaluating dLMs in the context of learners' age, special educational needs, and other context settings and the presented results are valuable for developing pre-service teachers' ability of „selecting dLM." Further, in teacher education, to develop this facet of digital competence for effective lesson planning, it is crucial to present dLMs as an option with their risks and affordances. Not knowing about dLMs and their affordances, pre-service teachers may not consider them in their later practices [48].

Similarly, responses regarding whether the dLM increases or decreases the extraneous cognitive load need to be examined with pre-service teachers in teacher education to ensure they know what and which additional scaffolding and support (sub-populations of) learners need when using dLMs [4]. Noteworthy is that none of the participants used the unavailability of digital technology required for using dLMs in classrooms in their reasoning, as done in other studies [48]. This could be because the wording of the items inferred it or the participants' understanding of the level of available technology in education in the local context. These are just some of the conclusions that can be drawn from the results, highlighting potential areas in which education institutions are effective and ineffective in their teaching—understanding the limitations of the results as we have previously outlined.

To assess teacher education effectiveness, it is equally vital that items can be adjusted to different teaching subjects and localized contexts—curriculum, school infrastructure, and local rules and regulations [44]. Our items and scoring have these affordances since the scoring of content knowledge (items 1-3) is evaluated in the context of the local curriculum and the teaching subject. Second, the scoring of the reasoning (item 4) for or against using the digital learning material was structured along TCK, TPK, and TPACK and can, for TPK, be directly transferred to other teaching subjects. The coding of TCK applies to the mathematics learning content and may require adjustments for the transfer to other subjects. Different and further reasons may apply. For example, the ability to simulate phenomena invisible to the eye in biology or chemistry [27,55], connecting physical phenomena with their representations in physics [52], or understanding content through audio cues in language education [56]. We suggest that future studies in this respect evaluate dLMs using the coding model shown in Table 1 and potentially contrast dLM with other (non-) dLM.

Also crucial for the longevity of assessment items used in teacher development is the ability to adjust the items and create variations so they can be reused without requiring retesting. An obvious modification is to use a different dLM. When doing so, one should use a dLM that is complex enough to provide various arguments [14,48]. Another option is, instead of inquiring about learner age and special education needs, they and other contextual factors can be given with a dLM, as well as the items and the requirement to argue for or against using a dLM with those confinements.

5.4. Limitations

Despite the theoretical support and empirical validation, our approach has several limitations. First, the empirical data is limited to mathematics pre-service teachers from two universities in Germany and Austria. The dLM used in the larger sample (study 2) assessed a simple mathematical concept primarily applicable to primary and special education pre-service teachers. The dLMs utilized in the initial study (study 1) were also limited to the learning content of primary education, mainly due to the sample population available to us. Second, methodologically, a limitation is that we did not evaluate outcome variables in teacher education [16], as neither institution offers specific courses on technology use for all pre-service teachers. Third, respondents volunteered to participate in the online test, and the data was not collected through an assessment. We believe that biases from

social desirability [36,cf. 37] played a minor role, and inattentive responses [97] were filtered out in study 2, but we cannot entirely dismiss these effects. Fourth, the composition of participants in both studies differed and had smaller subsets regarding the number of semesters and study stages, limiting the comparison of the two and the overall statistical analysis.

6. Conclusions

In summary, notwithstanding these limitations, we believe that our economic approach of assessing the facet of digital competence „selecting dLM," using only four items—two open and two closed items—provides a more focused and efficient approach compared to developing lesson plans [15,27] or the evaluation of developed lesson plans incorporating dRs [31,98,99], and especially via TPACK self-reports. In addition, the results of the developed four items represent an approach for the objective assessment of TPACK reasoning that enables other diagnostics, leading to other learning opportunities in teacher education than would be the case with self-reported TPACK.

In addition, our approach can be adapted to different teaching subjects, educational stages, and local contexts. We have validated the items already with pre-service teachers at two universities in different countries, thus proving the item's suitability for comparing the effectiveness of different teacher training systems [16] and especially university-based teacher training programs for varying levels of mathematics teaching (primary, secondary, and special education).

Appendix A

Appendix A.1

Table A1. Scoring examples for dLM1-dLM4 for items 1-3, learner age and special needs, and description of the learning content of the dLM.

score	dLM1 (circle, and property of radius) [76]	dLM2 (symmetry and axis of symmetry) [71]	dLM3 (generate bar charts) [72]	dLM4 (arithmetic mean) [73]
No, or wrong description (score 0)	„geometry, shapes, figures"	„spatial thinking"	„visualization, sizes and quantities "	„probability"

description inappropriate for learner age, or needs (score1)	7-8'th grade and no special educational needs; „circles and the properties of circles (radius…)"	7-8'th grade and no FSP„ symmetry "	1-2'nd grade and learning disabilities „bar charts"	7-8'th grade and hearing impairment „arithmetic mean"
generic description appropriate for learner age, or needs (score 2)	„the content is useful for introducing circles and their radius."	„axis mirroring"	„bar charts"	„arithmetic mean"
detailed description appropriate for learner age, or needs (score 3)	„To introduce the circle. Pupils should be made aware that every point on the circle is exactly the same distance from the center."	„the material presents that the dimensions of the mirrored object remain the same size when mirrored at a straight line."	„It is about absolute frequencies and the creation of bar charts."	„The dLM evaluates students' understanding of how to calculate the arithmetic mean, … helping students to grasp the underlying principles of the calculation."

Appendix A.2

Table A2. Scoring examples for dLM1-dLM4 for item 4, TCK, TPK, and TPACK reasoning.

score	dLM1 (circle, and property of radius) [76]	dLM2 (symmetry and axis of symmetry) [71]	dLM3 (generate bar charts) [72]	dLM4 (arithmetic mean) [73]
no, or wrong reasoning (score 0)	„ I don't see much point in the dLM"	„there are better visual examples "	„would rather do it analogue."	„I don't know"
generic reasoning (score1)	„a good concept that combines math with technology"	„good representation of the principle of symmetry"	„simpel und nice"	„a good task for calculating"
1 x TCK, or 1 x TPK reasoning (score 2)	„I wouldn't use the learning environment in a setting where students need support. It requires a lot of cognitive skills to comprehend the task and be able to visualize it…"	„It is fun and motivating for the children to watch how the butterfly can move its wings"	„It is a good activity to check whether students understand the representation of the bar chart without having to draw a chart themselves (saves time)."	„…learners can all work self-directed, there are solution hints…"
2 x TCK, or 2 x TPK reasoning (score 3)	„ Learners self-directed how it is possible to solve the task and thus learn an important property of the	„Students learn about symmetries through play, Students can learn about the properties of symmetries	„…motivational, context accessible to learners…"	„…as everyone can work on the tasks at their own pace and it can be a motivating factor for the children to

	circle (radius) in a playful way "	through experimentation, which would be more difficult without digital media"		work digitally and see results immediately..."
TCK and TPK, thus TPACK reasoning (score 4)	„ I would use this learning environment because it is enactive and visual learning that actively engages students in the learning process. Through the concrete task of positioning x in a circle around the tree, the children experience geometric concepts such as radius, center, and circle shape. This not only promotes an understanding of abstract mathematical concepts, but also spatial awareness and the ability to recognize connections"	„ I wouldn't use the learning environment... For example, the task is too abstract for learners or offer too few differentiated approaches to understand the core of axial symmetry. If there's no way to adapt the task to different learning levels, some students might be overwhelmed or under-challenged."	„...without requiring learners to do a lot of drawing. Learners can easily experiment and check their solutions independently. Doing this on paper would waste lesson time and verification of results is time consuming for teachers..."	„...It assesses students' understanding of calculating the arithmetic mean. Students are forced to rethink their learned knowledge of calculation and can thus better reflect on the arithmetic mean calculation. However, I view this learning environment more as a test to determine the extent to which students have internalized the subject matter they have learned. "

References

1. Drijvers, P.; Sinclair, N. The Role of Digital Technologies in Mathematics Education: Purposes and Perspectives. *ZDM Mathematics Education* **2023**, doi:10.1007/s11858-023-01535-x.

2. Roblyer, M.D.; Hughes, J.E. *Integrating Educational Technology into Teaching: Transforming Learning across Disciplines*; Eighth Edition.; Pearson Education, Inc: New York, 2019; ISBN 978-0-13-474641-8.

3. Bond, M.; Marín, V.I.; Dolch, C.; Bedenlier, S.; Zawacki-Richter, O. Digital Transformation in German Higher Education: Student and Teacher Perceptions and Usage of Digital Media. *Int J Educ Technol High Educ* **2018**, *15*, 48, doi:10.1186/s41239-018-0130-1.

4. Weinhandl, R.; Houghton, T.; Lindenbauer, E.; Mayerhofer, M.; Lavicza, Z.; Hohenwarter, M. Integrating Technologies Into Teaching and Learning Mathematics at the Beginning of Secondary Education in Austria. *EURASIA J Math Sci Tech Ed* **2021**, *17*, 1–15, doi:10.29333/ejmste/11428.

5. Borba, M.C. The Future of Mathematics Education since COVID-19: Humans-with-Media or Humans-with-Non-Living-Things. *Educ Stud Math* **2021**, *108*, 385–400, doi:10.1007/s10649-021-10043-2.

6.   Haleem, A.; Javaid, M.; Qadri, M.A.; Suman, R. Understanding the Role of Digital Technologies in Education: A Review. *Sustainable Operations and Computers* **2022**, *3*, 275–285, doi:10.1016/j.susoc.2022.05.004.

7.   Huber, S.G.; Helm, C. COVID-19 and Schooling: Evaluation, Assessment and Accountability in Times of Crises—Reacting Quickly to Explore Key Issues for Policy, Practice and Research with the School Barometer. *Educ Asse Eval Acc* **2020**, *32*, 237–270, doi:10.1007/s11092-020-09322-y.

8.   Kaspar, K.; Burtniak, K.; Rüth, M. Online Learning during the Covid-19 Pandemic: How University Students' Perceptions, Engagement, and Performance Are Related to Their Personal Characteristics. *Curr Psychol* **2024**, *43*, 16711–16730, doi:10.1007/s12144-023-04403-9.

9.   Engelbrecht, J.; Llinares, S.; Borba, M.C. Transformation of the Mathematics Classroom with the Internet. *ZDM Mathematics Education* **2020**, *52*, 825–841, doi:10.1007/s11858-020-01176-4.

10.   OECD *The Future of Education and Skills Education 2030*; OECD: Paris, France, 2018;

11.   Heine, S.; König, J.; Krepf, M. Digital Resources as an Aspect of Teacher Professional Digital Competence: One Term, Different Definitions – a Systematic Review. *Education and Information Technologies* **2022**, 28, doi:https://doi.org/10.1007/s10639-022-11321-z.

12.   Weigand, H.-G.; Trgalova, J.; Tabach, M. Mathematics Teaching, Learning, and Assessment in the Digital Age. *ZDM Mathematics Education* **2024**, doi:10.1007/s11858-024-01612-9.

13.   Clark-Wilson, A.; Robutti, O.; Thomas, M. Teaching with Digital Technology. *ZDM Mathematics Education* **2020**, *52*, 1223–1242, doi:10.1007/s11858-020-01196-0.

14.   Lindenbauer, E.; Infanger, E.-M.; Lavicza, Z. Enhancing Mathematics Education through Collaborative Digital Material Design: Lessons from a National Project. *EUR J SCI MATH ED* **2024**, *12*, 276–296, doi:10.30935/scimath/14323.

15.   Gonscherowski, P.; Rott, B. A Systematic Review of the Literature on TPACK Instruments Used with Pre-Service Teachers from 2017 to 2023 Focused on Selecting Digital Resources. *Journal of Computers in Education* **Accepted**.

16.   König, J.; Heine, S.; Kramer, C.; Weyers, J.; Becker-Mrotzek, M.; Großschedl, J.; Hanisch, C.; Hanke, P.; Hennemann, T.; Jost, J.; et al. Teacher Education Effectiveness as an Emerging Research Paradigm: A Synthesis of Reviews of Empirical Studies Published over Three Decades (1993–2023). *Journal of Curriculum Studies* **2023**, 1–21, doi:10.1080/00220272.2023.2268702.

17.   Schmidt, W.H.; Xin, T.; Guo, S.; Wang, X. Achieving Excellence and Equality in Mathematics: Two Degrees of Freedom? *Journal of Curriculum Studies* **2022**, *54*, 772–791, doi:10.1080/00220272.2022.2098536.

18.   Reinhold, F.; Leuders, T.; Loibl, K.; Nückles, M.; Beege, M.; Boelmann, J.M. Learning Mechanisms Explaining Learning With Digital Tools in Educational Settings: A Cognitive Process Framework. *Educ Psychol Rev* **2024**, *36*, 14, doi:10.1007/s10648-024-09845-6.

19.   Anderson, S.; Griffith, R.; Crawford, L. TPACK in Special Education: Preservice Teacher Decision Making While Integrating iPads into Instruction. *Contemporary Issues in Technology and Teacher Education (CITE Journal)* **2017**, *17*.

20.   Tseng, S.-S.; Yeh, H.-C. Fostering EFL Teachers' CALL Competencies through Project-Based Learning. *Educational Technology & Society* **2019**, *22*, 94–105.

21.   Handal, B.; Campbell, C.; Cavanagh, M.; Petocz, P. Characterising the Perceived Value of Mathematics Educational Apps in Preservice Teachers. *Mathematics Education Research Journal* **2022**, 1–23.

22. Valtonen, T.; Leppänen, U.; Hyypiä, M.; Sointu, E.; Smits, A.; Tondeur, J. Fresh Perspectives on TPACK: Pre-Service Teachers' Own Appraisal of Their Challenging and Confident TPACK Areas. *Educ Inf Technol* **2020**, *25*, 2823–2842, doi:10.1007/s10639-019-10092-4.

23. Mishra, P.; Koehler, M.J. Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge. *Teachers College Record* **2005**, *108*, 1017–1054, doi:10.1111/j.1467-9620.2006.00684.x.

24. Koehler, M.J.; Shin, T.S.; Mishra, P. How Do We Measure TPACK? Let Me Count the Ways. In *A Research Handbook on Frameworks and Approaches*; Ronau, R.N., Rakes, C.R., Niess, M.L., Eds.; IGI Global, 2011; pp. 16–31 ISBN 978-1-60960-750-0.

25. Schmid, M.; Brianza, E.; Mok, S.Y.; Petko, D. Running in Circles: A Systematic Review of Reviews on Technological Pedagogical Content Knowledge (TPACK). *Computers & Education* **2024**, *214*, 105024, doi:10.1016/j.compedu.2024.105024.

26. Karakaya Cirit, D.; Canpolat, E. A Study on the Technological Pedagogical Contextual Knowledge of Science Teacher Candidates across Different Years of Study. *Education and Information Technologies* **2019**, *24*, 2283–2309.

27. von Kotzebue, L. Beliefs, Self-Reported or Performance-Assessed TPACK: What Can Predict the Quality of Technology-Enhanced Biology Lesson Plans? *Journal of Science Education and Technology* **2022**, *31*, 570–582.

28. Lachner, A.; Fabian, A.; Franke, U.; Preiß, J.; Jacob, L.; Führer, C.; Küchler, U.; Paravicini, W.; Randler, C.; Thomas, P. Fostering Pre-Service Teachers' Technological Pedagogical Content Knowledge (TPACK): A Quasi-Experimental Field Study. *Computers & Education* **2021**, *174*.

29. Jin, Y.; Harp, C. Examining Preservice Teachers' TPACK, Attitudes, Self-Efficacy, and Perceptions of Teamwork in a Stand-Alone Educational Technology Course Using Flipped Classroom or Flipped Team-Based Learning Pedagogies. *Journal of Digital Learning in Teacher Education* **2020**, *36*, 166–184.

30. Mouza, C.; Nandakumar, R.; Yilmaz Ozden, S.; Karchmer-Klein, R. A Longitudinal Examination of Preservice Teachers' Technological Pedagogical Content Knowledge in the Context of Undergraduate Teacher Education. *Action in Teacher Education* **2017**, *39*, 153–171.

31. Pekkan, Z.T.; Ünal, G. Technology Use: Analysis of Lesson Plans on Fractions in an Online Laboratory School. In *Proceedings of the 45th Conference of the International Group for the Psychology of Mathematics Education*; Fernández, C., Llinares, S., Gutiérrez, A., Planas, N., Eds.; Alicante/Spain, 2022; Vol. 4, p. 410 ISBN 978-84-1302-178-2.

32. Gonscherowski, P.; Rott, B. Selecting Digital Technology: A Review of TPACK Instruments. In *Proceedings of the 46th conference of the International Group for the Psychology of Mathematics Education*; Ayalon, M., Koichu, B., Leikin, R., Rubel, L., Tabach, M., Eds.; PME: Haifa, Israel, 2023; Vol. 2, pp. 378–386.

33. McCulloch, A.; Leatham, K.; Bailey, N.; Cayton, C.; Fye, K.; Lovett, J. Theoretically Framing the Pedagogy of Learning to Teach Mathematics with Technology. *Contemporary Issues in Technology and Teacher Education (CITE Journal)* **2021**, *21*.

34. Revuelta-Domínguez, F.-I.; Guerra-Antequera, J.; González-Pérez, A.; Pedrera-Rodríguez, M.-I.; González-Fernández, A. Digital Teaching Competence: A Systematic Review. *Sustainability* **2022**, *14*, 6428, doi:10.3390/su14116428.

35. Yeh, Y.; Hsu, Y.; Wu, H.; Hwang, F.; Lin, T. Developing and Validating Technological Pedagogical Content Knowledge-practical TPACK through the Delphi Survey Technique. *Brit J Educational Tech* **2014**, *45*, 707–722, doi:10.1111/bjet.12078.

36. Grimm, P. Social Desirability Bias. In *Wiley International Encyclopedia of Marketing*; Sheth, J., Malhotra, N., Eds.; Wiley, 2010 ISBN 978-1-4051-6178-7.

37. Safrudiannur Measuring Teachers' Beliefs Quantitatively: Criticizing the Use of Likert Scale and Offering a New Approach, Springer Fachmedien Wiesbaden: Wiesbaden, 2020.

38. Schmid, M.; Brianza, E.; Petko, D. Self-Reported Technological Pedagogical Content Knowledge (TPACK) of Pre-Service Teachers in Relation to Digital Technology Use in Lesson Plans. *Computers in Human Behavior* **2021**, *115*, 106586, doi:10.1016/j.chb.2020.106586.

39. Blömeke, S.; Gustafsson, J.-E.; Shavelson, R.J. Beyond Dichotomies: Competence Viewed as a Continuum. *Zeitschrift für Psychologie* **2015**, *223*, 3–13, doi:10.1027/2151-2604/a000194.

40. Deng, Z. Powerful Knowledge, Educational Potential and Knowledge-Rich Curriculum: Pushing the Boundaries. *Journal of Curriculum Studies* **2022**, *54*, 599–617, doi:10.1080/00220272.2022.2089538.

41. Yang, X.; Deng, J.; Sun, X.; Kaiser, G. The Relationship between Opportunities to Learn in Teacher Education and Chinese Preservice Teachers' Professional Competence. *Journal of Curriculum Studies* **2024**, 1–19, doi:10.1080/00220272.2024.2334944.

42. Schleicher, A. *PISA 2022 Insights and Interpretations*; OECD, 2023;

43. Kaiser, G.; König, J. Competence Measurement in (Mathematics) Teacher Education and Beyond: Implications for Policy. *High Educ Policy* **2019**, *32*, 597–615, doi:10.1057/s41307-019-00139-z.

44. Koehler, M.J.; Mishra, P.; Cain, W. What Is Technological Pedagogical Content Knowledge (TPACK)? *Journal of Education* **2013**, *193*, 13–19, doi:10.1177/002205741319300303.

45. Shulman, L.S. Those Who Understand: Knowledge Growth in Teaching. *Educational Researcher* **1986**, *15*, 4–14, doi:10.3102/0013189X015002004.

46. Mishra, P.; Warr, M. Contextualizing TPACK within Systems and Cultures of Practice. *Computers in Human Behavior* **2021**, *117*, 106673, doi:10.1016/j.chb.2020.106673.

47. Tabach, M.; Trgalová, J. Teaching Mathematics in the Digital Era: Standards and Beyond. In *STEM Teachers and Teaching in the Digital Era*; Ben-David Kolikant, Y., Martinovic, D., Milner-Bolotin, M., Eds.; Springer International Publishing: Cham, Switzerland, 2020; pp. 221–242 ISBN 978-3-030-29395-6.

48. Gonscherowski, P.; Rott, B. How Do Pre-/In-Service Mathematics Teachers Reason for or against the Use of Digital Technology in Teaching? *Mathematics* **2022**, *10*, 2345, doi:10.3390/math10132345.

49. Bonafini, F.C.; Lee, Y. Investigating Prospective Teachers' TPACK and Their Use of Mathematical Action Technologies as They Create Screencast Video Lessons on iPads. *TechTrends: Linking Research and Practice to Improve Learning* **2021**, *65*, 303–319.

50. Hillmayr, D.; Ziernwald, L.; Reinhold, F.; Hofer, S.I.; Reiss, K.M. The Potential of Digital Tools to Enhance Mathematics and Science Learning in Secondary Schools: A Context-Specific Meta-Analysis. *Computers & Education* **2020**, *153*, 103897, doi:10.1016/j.compedu.2020.103897.

51. Morgan, C.; Kynigos, C. Digital Artefacts as Representations: Forging Connections between a Constructionist and a Social Semiotic Perspective. *Educ Stud Math* **2014**, *85*, 357–379, doi:10.1007/s10649-013-9523-1.

52. Solvang, L.; Haglund, J. How Can GeoGebra Support Physics Education in Upper-Secondary School -- A Review. *Physics Education* **2021**, *56*.

53. Moreno-Armella, L.; Hegedus, S.J.; Kaput, J.J. From Static to Dynamic Mathematics: Historical and Representational Perspectives. *Educ Stud Math* **2008**, *68*, 99–111, doi:10.1007/s10649-008-9116-6.

54. Sweller, J.; van Merriënboer, J.J.G.; Paas, F. Cognitive Architecture and Instructional Design: 20 Years Later. *Educ Psychol Rev* **2019**, *31*, 261–292, doi:10.1007/s10648-019-09465-5.

55. Puspitasari, J.R.; Yamtinah, S.; Susilowati, E.; Kristyasari, M.L. Validation of TTMC Instrument of Pre-Service Chemistry Teacher's TPACK Using Rasch Model Application. *Journal of Physics: Conference Series* **2020**, *1511*.

56. Tseng, S.-S.; Yeh, H.-C. Fostering EFL Teachers' CALL Competencies through Project-Based Learning. *Educational Technology & Society* **2019**, *22*, 94–105.

57. Gerhard, K.; Jäger-Biela, D.J.; König, J. Opportunities to Learn, Technological Pedagogical Knowledge, and Personal Factors of Pre-Service Teachers: Understanding the Link between Teacher Education Program Characteristics and Student Teacher Learning Outcomes in Times of Digitalization. *Z Erziehungswiss* **2023**, *26*, 653–676, doi:10.1007/s11618-023-01162-y.

58. Molenaar, I.; Boxtel, C.; Sleegers, P. Metacognitive Scaffolding in an Innovative Learning Arrangement. *Instructional Science* **2011**, *39*, 785–803, doi:10.1007/s11251-010-9154-1.

59. Turan, Z.; Karabey, S.C. The Use of Immersive Technologies in Distance Education: A Systematic Review. *Educ Inf Technol* **2023**, doi:10.1007/s10639-023-11849-8.

60. Rüth, M.; Breuer, J.; Zimmermann, D.; Kaspar, K. The Effects of Different Feedback Types on Learning With Mobile Quiz Apps. *Front. Psychol.* **2021**, *12*, 665144, doi:10.3389/fpsyg.2021.665144.

61. Grossman, P.L. *Teaching Core Practices in Teacher Education*; Core Practices in Education Series; Harvard Education Press: Cambridge, Massachusetts, 2018; ISBN 978-1-68253-187-7.

62. Anderson, S.; Putman, R. Special Education Teachers' Experience, Confidence, Beliefs, and Knowledge about Integrating Technology. *Journal of Special Education Technology* **2020**, *35*, 37–50.

63. Redecker, C.; Punie, Y. European Framework for the Digital Competence of Educators – DigCompEdu. *Publications Office of the European Union* **2017**, doi:10.2760/159770.

64. Bonafini, F.C.; Lee, Y. Investigating Prospective Teachers' TPACK and Their Use of Mathematical Action Technologies as They Create Screencast Video Lessons on iPads. *TechTrends* **2021**, *65*, 303–319.

65. Drijvers, P. Digital Technology in Mathematics Education: Why It Works (Or Doesn't). In *Selected Regular Lectures from the 12th International Congress on Mathematical Education*; Cho, S.J., Ed.; Springer International Publishing: Cham, 2015; pp. 135–151 ISBN 978-3-319-17186-9.

66. Drijvers, P.; Ball, L.; Barzel, B.; Heid, M.K.; Cao, Y.; Maschietto, M. *Uses of Technology in Lower Secondary Mathematics Education: A Concise Topical Survey*; Kaiser, G., Ed.; ICME-13 Topical Surveys; Springer International Publishing: Cham, 2016; ISBN 978-3-319-33665-7.

67. Gonscherowski, P.; Rott, B. Measuring Digital Competencies of Pre-Service Teachers-a Pilot Study. In Proceedings of the Proceedings of the 44th Conference of the International Group for the Psychology of Mathematics Education; Inprasitha, M., Changsri, N., Boonsena, N., Eds.; Khon Kaen, Thailand, 2021; Vol. 1, p. 143.

68. Gonscherowski, P.; Rott, B. Digital Competencies of Pre-/in-Service Teachers-an Interview Study. In Proceedings of the Proceedings of the Twelfth Congress of the European Society for Research in Mathematics Education; Hodgen, J., Geraniou, E., Ferretti, F., Eds.; Free University of Bozen-Bolzano and ERME, 2022; pp. 2544–2551.

69. Gonscherowski, P.; Rott, B. Instrument to Assess the Knowledge and the Skills of Mathematics Educators' Regarding Digital Technology. In Proceedings of the Beiträge zum Mathematikunterricht 2022; WTM: Frankfurt/Germany, 2022; Vol. 3, p. 1424.

70. GeoGebra Team Classroom Resources Available online: https://www.geogebra.org/materials (accessed on 23 May 2023).

71. Schüngel, M. Bewege Den Schmetterling – GeoGebra Available online: https://www.geogebra.org/m/zrj2zcam (accessed on 28 February 2022).

72. FLINK Lieblingssport – GeoGebra Available online: https://www.geogebra.org/m/v4xuvmhf (accessed on 28 February 2023).

73. FLINK Welche Zahl Fehlt? – GeoGebra Available online: https://www.geogebra.org/m/qqv3kxt6 (accessed on 28 February 2023).

74. BMBF Lehrplan der allgemeinen Sonderschule, Available online: https://www.ris.bka.gv.at/Dokumente/BgblAuth/BGBLA_2008_II_137/COO_2026_100_2_440355.html (accessed on 28 August 2023).

75. MSB NRW Schulentwicklung NRW - Vorgaben Zieldifferente Bildungsgänge Available online: https://www.schulentwicklung.nrw.de/lehrplaene/vorgaben-zieldifferente-bildungsgaenge/index.html (accessed on 3 September 2023).

76. Flink Maxi und der Baum – GeoGebra Available online: https://www.geogebra.org/m/a4pppe7a (accessed on 10 July 2021).

77. Cohen, J. Quantitative Methods In Psychology: A Power Primer. *Psychological Bulletin* **1992**, *112*, 155–159.

78. Cicchetti, D.V. Guidelines, Criteria, and Rules of Thumb for Evaluating Normed and Standardized Assessment Instruments in Psychology. *Psychological Assessment* **1994**, *6*, 284–290, doi:10.1037/1040-3590.6.4.284.

79. Guillén-Gámez, F.D.; Mayorga-Fernández, M.J.; Bravo-Agapito, J.; Escribano-Ortiz, D. Analysis of Teachers' Pedagogical Digital Competence: Identification of Factors Predicting Their Acquisition. *Tech Know Learn* **2021**, *26*, 481–498, doi:10.1007/s10758-019-09432-7.

80. Rott, B. Inductive and Deductive Justification of Knowledge: Epistemological Beliefs and Critical Thinking at the Beginning of Studying Mathematics. *Educ Stud Math* **2021**, *106*, 117–132, doi:10.1007/s10649-020-10004-1.

81. Landis, J.R.; Koch, G.G. The Measurement of Observer Agreement for Categorical Data. *Biometrics* **1977**, *33*, 159–174, doi:10.2307/2529310.

82. Koh, K. Univariate Normal Distribution. In *Encyclopedia of Quality of Life and Well-Being Research*; Michalos, A.C., Ed.; Springer Netherlands: Dordrecht, 2014; pp. 6817–6819 ISBN 978-94-007-0752-8.

83. Taber, K.S. The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Res Sci Educ* **2018**, *48*, 1273–1296, doi:10.1007/s11165-016-9602-2.

84. Janssen, N.; Knoef, M.; Lazonder, A.W. Technological and Pedagogical Support for Pre-Service Teachers' Lesson Planning. *Technology, Pedagogy and Education* **2019**, *28*, 115–128.

85. Knievel, I.; Lindmeier, A.M.; Heinze, A. Beyond Knowledge: Measuring Primary Teachers' Subject-Specific Competences in and for Teaching Mathematics with Items Based on Video Vignettes. *Int J of Sci and Math Educ* **2015**, *13*, 309–329, doi:10.1007/s10763-014-9608-z.

86. Weyers, J.; Kramer, C.; Kaspar, K.; König, J. Measuring Pre-Service Teachers' Decision-Making in Classroom Management: A Video-Based Assessment Approach. *Teaching and Teacher Education* **2024**, *138*, 104426, doi:10.1016/j.tate.2023.104426.

87. Baumanns, L.; Pohl, M. Leveraging ChatGPT for Problem Posing: An Exploratory Study of Pre-Service Teachers' Professional Use of AI. **2024**.

88. Cai, J.; Rott, B. On Understanding Mathematical Problem-Posing Processes. *ZDM Mathematics Education* **2024**, *56*, 61–71, doi:10.1007/s11858-023-01536-w.

89. Trixa, J.; Kaspar, K. Information Literacy in the Digital Age: Information Sources, Evaluation Strategies, and Perceived Teaching Competences of Pre-Service Teachers. *Front. Psychol.* **2024**, *15*, 1336436, doi:10.3389/fpsyg.2024.1336436.

90. Seifert, H.; Kosiol, T.; Gonscherowski, P.; Rott, B.; Ufer, S.; Lindmeier, A. Using a Futures Study Methodology to Explore the Impact of New Technologies on Mathematics Teachers' Core Practices and Professional Knowledge In Review.

91. König, J.; Doll, J.; Buchholtz, N.; Förster, S.; Kaspar, K.; Rühl, A.-M.; Strauß, S.; Bremerich-Vos, A.; Fladung, I.; Kaiser, G. Pädagogisches Wissen versus fachdidaktisches Wissen?: Struktur des professionellen Wissens bei angehenden Deutsch-, Englisch- und Mathematiklehrkräften im Studium. *Z Erziehungswiss* **2018**, *21*, 1–38, doi:10.1007/s11618-017-0765-z.

92. König, J.; Heine, S.; Jäger-Biela, D.; Rothland, M. ICT Integration in Teachers' Lesson Plans: A Scoping Review of Empirical Studies. *European Journal of Teacher Education* **2022**, 1–29, doi:10.1080/02619768.2022.2138323.

93. Thurm, D.; Barzel, B. Teaching Mathematics with Technology: A Multidimensional Analysis of Teacher Beliefs. *Educ Stud Math* **2022**, *109*, 41–63, doi:10.1007/s10649-021-10072-x.

94. Chen, W.; Tan, J.S.H.; Pi, Z. The Spiral Model of Collaborative Knowledge Improvement: An Exploratory Study of a Networked Collaborative Classroom. *International Journal of Computer-Supported Collaborative Learning* **2021**, *16*, 7–35.

95. Karakaya Cirit, D.; Canpolat, E. A Study on the Technological Pedagogical Contextual Knowledge of Science Teacher Candidates across Different Years of Study. *Education and Information Technologies* **2019**, *24*, 2283–2309.

96. Purwaningsih, E.; Nurhadi, D.; Masjkur, K. TPACK Development of Prospective Physics Teachers to Ease the Achievement of Learning Objectives: A Case Study at the State University of Malang, Indonesia. *Journal of Physics: Conference Series* **2019**, *1185*.

97. Curran, P.G. Methods for the Detection of Carelessly Invalid Responses in Survey Data. *Journal of Experimental Social Psychology* **2016**, *66*, 4–19, doi:10.1016/j.jesp.2015.07.006.

98. Jin, Y.; Harp, C. Examining Preservice Teachers' TPACK, Attitudes, Self-Efficacy, and Perceptions of Teamwork in a Stand-Alone Educational Technology Course Using Flipped Classroom or Flipped Team-Based Learning Pedagogies. *Journal of Digital Learning in Teacher Education* **2020**, *36*, 166–184.

99. Mouza, C.; Nandakumar, R.; Yilmaz Ozden, S.; Karchmer-Klein, R. A Longitudinal Examination of Preservice Teachers' Technological Pedagogical Content Knowledge in the Context of Undergraduate Teacher Education. *Action in Teacher Education* **2017**, *39*, 153–171.