



OPEN ACCESS

EDITED BY

Till Bruckermann,
Leibniz University Hannover, Germany

REVIEWED BY

Lena von Kotzebue,
University of Salzburg, Austria
Alexander Georg Buessing,
Leibniz University Hannover, Germany

*CORRESPONDENCE

Alexander Aumann
✉ alexander.aumann@ph-weingarten.de

RECEIVED 20 March 2023

ACCEPTED 11 September 2023

PUBLISHED 29 September 2023

CITATION

Aumann A, Schnebel S and Weitzel H (2023)
The EnTPACK rubric: development, validation,
and reliability of an instrument for measuring
pre-service science teachers' enacted TPACK.
Front. Educ. 8:1190152.
doi: 10.3389/educ.2023.1190152

COPYRIGHT

© 2023 Aumann, Schnebel and Weitzel. This is
an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with
these terms.

The EnTPACK rubric: development, validation, and reliability of an instrument for measuring pre-service science teachers' enacted TPACK

Alexander Aumann^{1*}, Stefanie Schnebel² and Holger Weitzel¹

¹Department of Biology, University of Education Weingarten, Weingarten, Germany, ²Department of Educational Science, University of Education Weingarten, Weingarten, Germany

Science instruction can benefit from the use of digital technologies if pre-service teachers are given opportunities to acquire Technological Pedagogical And Content Knowledge (TPACK) as part of their studies. However, the prevailing self-report approach to TPACK measurement does not allow conclusions to be drawn about enacted TPACK, which is rarely assessed in real classroom situations. In addition, instruments designed to measure TPACK enactment lack descriptive clarity and no single instrument is used to assess the three relevant phases of teacher competencies (lesson planning, implementation, and reflection). The present paper addresses this gap by presenting the development and validation of a comprehensive rubric for assessing the enacted TPACK of pre-service science teachers. To operationalize the “fuzzy” aspects of the framework, the rubric targets a specific use of digital media and instructional approach in science teaching: student-generated explainer videos and animations. At the core of the development process is a theory- and literature-based systematic review of (1) existing instruments for assessing pre-service science teachers' enacted TPACK and (2) instructional criteria for student-generated explainer videos in science classes. The resulting rubric allows valid conclusions given the appropriate conditions, has demonstrated reliability, and excels due to its specific focus, high degree of differentiation, systematic grounding in theory and literature, objective grading criteria, and comprehensive applicability to all three phases of teacher competencies.

KEYWORDS

enacted TPACK, rubric development, student-generated animation, systematic review, teacher professionalization, student-created explainer videos, science teaching

1. Introduction

The ongoing digital revolution has transformed society, resulting in educational opportunities and challenges (KMK, 2017; Redecker, 2017). Among other benefits, the availability of digital technology and media (DTM) at a relatively low cost facilitates the visualization of abstract and complex scientific content (Hsu et al., 2015) and enables students to engage in increased interactive learning (Chi and Wylie, 2014). While the benefits

of using DTM in science teaching are obvious, simply incorporating DTM into the classroom does not automatically result in high-quality teaching (Kates et al., 2018). For teachers to harness the full potential of DTM, they must be purposefully integrated into instructional practices and aligned with learning objectives (Petko, 2020). To achieve this, pre-service science teachers (PSSTs) need to acquire a complex body of Technological Pedagogical And Content Knowledge (TPACK; Mishra and Koehler, 2006) as part of their studies (Becker et al., 2020). The TPACK model extends Shulman's (1986) Pedagogical Content Knowledge by the fundamental knowledge facet of Technological Knowledge, resulting in a total of seven relevant knowledge facets: Technological Knowledge (TK), Pedagogical Knowledge (PK), Content Knowledge (CK), Technological Content Knowledge (TCK), Technological Pedagogical Knowledge (TPK), Pedagogical Content Knowledge (PCK), and lastly TPACK. Although the scientific conceptualization of teacher competencies with regard to DTM is currently extended to Digitally-Related Pedagogical And Content Knowledge (DPACK; Thyssen et al., 2023), this study addresses TPACK as an appropriate model for two reasons: (1) The DPACK model links digital competence with digital literacy by adding sociocultural knowledge as a fourth fundamental knowledge facet. Although this represents a reasonable extension of the TPACK model in terms of a comprehensive description of digitization-related teacher competencies, this adaptation results in a increased model complexity. Since the present study focuses on digital competencies of PSSTs, an increased complexity due to the inclusion of sociocultural aspects is avoided. (2) There is already a considerable body of empirical evidence on the topic of TPACK in science teacher education that the study can draw on.

1.1. Professional development of teachers

There has been a growing body of evidence that shows how TPACK could be acquired in teacher education (Tondeur et al., 2012, 2018). Nevertheless, how to transfer professional knowledge about teaching into classroom practice remains a fundamental question within science education (Labudde and Möller, 2012) and also relates to the successful integration of DTM (Pareto and Willermark, 2019). In the transformation model of lesson planning (Stender et al., 2017), professional development occurs when professional knowledge acquired at university is transformed into the knowledge of experienced teachers, where it is represented as highly automated teaching scripts. This transformative process depends on repeated lesson planning, implementation, and reflection on teaching practice. Concepts of teacher professional development indicate that the three phases of planning, acting, and reflecting are closely interrelated and necessary for the successful development of teaching competencies (e.g., Baumert and Kunter, 2013; Stender et al., 2017).

1.2. Enacted TPACK

Different models have been developed to represent TPACK enactment from different perspectives, such as TPACK-Practical

(Hsu et al., 2015), TPACK *in situ* (Pareto and Willermark, 2019), and TPACK-in-action (Ling Koh et al., 2014). To collectively address the main content of these different models, we use the term *enacted TPACK* in this paper. In line with the transformation model of lesson planning and concepts of teacher professional development, enacted TPACK refers to technological, pedagogical, and content knowledge that is expanded over time within teaching situations (Ay et al., 2015; Pareto and Willermark, 2019). Thus, teaching experience is a constitutive part of TPACK enactment (Ay et al., 2015; Yeh et al., 2015; Jen et al., 2016), acting as both a driving resource and influence on TPACK (Hsu et al., 2015) and further framed by situationally unique contextual factors (Ling Koh et al., 2014; Pareto and Willermark, 2019; Brianza et al., 2022). Accordingly, enacted TPACK can be interpreted as a dynamic, contextual, and situated body of practical knowledge that guides lesson planning, implementation, and reflection (Hsu et al., 2015; Willermark, 2018; Pareto and Willermark, 2019).

1.3. TPACK measurement

Studies assessing TPACK are largely based on self-report measures (Voogt et al., 2012; Wang et al., 2018; Willermark, 2018), requiring participants to evaluate their knowledge of the seven TPACK facets—mostly without subject-specific contextualization (Chai et al., 2016). However, the quality of self-report measurement instruments frequently suffers from the insufficient linkage between knowledge that is self-reported and knowledge assessed by performance tests (e.g., Drummond and Sweeney, 2017), and self-report findings are hardly correlated at all with the enactment of TPACK (So and Kim, 2009; Mourlam et al., 2021). In this regard, the accuracy of PSSTs self-assessment seems to decrease as the amount of procedural (vs. declarative) knowledge required to complete tasks increases (Max et al., 2022). Consequently, self-reports should be supplemented with external assessment methods to increase objectivity (Max et al., 2022). Some approaches use knowledge tests such as those requiring participants to identify true or false TPACK-related statements (Drummond and Sweeney, 2017), or teaching vignettes in which PSSTs describe actions in a TPACK-related classroom situation (Max et al., 2022). Such assessments are closer to teaching practice but reduce the complexity of the teaching-learning process since moderating factors such as motivation or beliefs are not considered (Baumert and Kunter, 2013; Stender et al., 2017; Huwer et al., 2019). This also applies to studies using microteaching situations for TPACK measurement (e.g., Yeh et al., 2015; Canbazoglu Bilici et al., 2016; Jen et al., 2016; Aktaş and Özmen, 2020). To fully account for the contextual and situated nature of enacted TPACK, real classroom assessment must occur (Ay et al., 2015; Rosenberg and Koehler, 2015; Willermark, 2018; Pareto and Willermark, 2019). Furthermore, to accurately measure all facets of the construct, research instruments should be specific in terms of technology (DTM), pedagogy (instructional approach), and content (Njiku et al., 2020). Several reviews point to a lack of TPACK studies focusing on specific subject domains (Voogt et al., 2012; Wu, 2013; Willermark, 2018) while other researchers highlight a corresponding lack of subject-specific measurement tools (von Kotzebue, 2022).

1.4. Purpose of the study

The current study aimed to develop and validate an instrument to comprehensively measure TPACK as enacted by PSSTs. Taking into account the Transformation Model of Lesson Planning (Stender et al., 2017) as well as corresponding recommendations regarding enacted TPACK research (Willermark, 2018; Mourlam et al., 2021), we intended to develop an instrument that permits triangulated and comparable assessment of planning, implementation, and reflection on science teaching. Given the lack of clear operationalizations and descriptions of instruments currently used to assess enacted TPACK in the field (Willermark, 2018), the instrument was tailored to a specific use of DTM and a corresponding instructional approach. Attention was paid to selecting a DTM and instructional approach that (1) could offer a high potential in science teaching, (2) would be technologically feasible in school conditions, and (3) could be used flexibly with a variety of science topics.

In this regard, we concentrate on student-generated explainer videos or animations (SGEVA). This use of DTM allows students to actively engage with scientific content by transferring it into a new form of representation (video or animation). In this process, students interact with peers in discursive conversations about the subject content, promoting the expression of technical language. Hence, there are several reasons why the instrument focuses on SGEVA in science classes. First, there is growing evidence for the effectiveness of SGEVA in learning science concepts (Hoban and Nielsen, 2010; Hoban, 2020), as students become active and productive players in the classroom, stimulating a more in-depth engagement with the science content (Gallardo-Williams et al., 2020; Hoban, 2020). Second, and related to its technological intuitivity and economic feasibility, the use of SGEVA is steadily increasing in science classrooms (Gallardo-Williams et al., 2020; Potter et al., 2021). Third, by combining the rather generic competence areas of “presentation” and “communication and collaboration,” SGEVA can be used in various topics or domains compared to highly science-specific areas such as “measurement and data acquisition” (Becker et al., 2020). Given the shortage of both clearly operationalized instruments assessing the enactment of TPACK in the field and well-defined criteria of instructional quality using SGEVA in science classrooms, the enacted TPACK instrument (EnTPACK rubric) was developed with the following objectives in mind:

- (1) Analyzing instruments for assessing enacted TPACK using lesson plans, lesson observations, and/or lesson reflections.
- (2) Defining criteria for the instructional quality of a science lesson that involves students creating explainer videos or animations.
- (3) Developing and validating an instrument for assessing a specific aspect of PSSTs enacted TPACK using lesson plans, observations, and reflections.

2. Materials and methods

The development of the instrument was oriented to the major phases of analysis, design, development, implementation, and

evaluation (according to the ADDIE model), which are widely adopted in instructional design (Molenda, 2015).

- *Analysis* of the rubric’s objective in terms of the contextual conditions and definition of the theoretical basis.
- *Design*: Addressing the objectives by two systematic reviews of the literature and manifesting the theoretical basis.
- *Development* and operationalization of an initial instrument.
- *Implementation* of the instrument in field trials.
- *Evaluation* of content validity and reliability.

2.1. Analysis: determining the rubric’s objective and defining the theoretical basis

Initially, the objective of the rubric and the theoretical basis were defined. For this purpose, the rubric’s prospective application context was determined and appropriate theoretical models were selected. Taking the TPACK model as a starting point, the rubric was specified on the knowledge facets TK, PK, and CK as the core dimensions of the model. Building on the core dimensions, TPACK facets were identified to allow comprehensive insights into the specific application context. These facets were then theoretically substantiated by the incorporation of selected theoretical models.

2.2. Design: conducting the literature reviews

The theoretical foundation was empirically substantiated by incorporating current research literature from two areas: first, instruments for measuring the enacted TPACK of PSSTs, and second, instructional criteria for designing lessons that integrate SGEVA into science classes. To this end, we conducted a systematic review for both areas. This includes specification regarding the main objectives of the reviews, the selected information sources, the eligibility criteria, and the included studies as well as a summary of the results relevant to the objectives.

Both reviews utilized Google Scholar, which provides the most comprehensive overview of literature compared to 12 of the most widely consulted academic search engines and databases such as Web of Science or Scopus (Gusenbauer, 2019). Since it is not possible to conduct complex search queries in the Google Scholar web version, title searches were performed via Harzing’s “Publish or Perish” software (version 8). The final searches were conducted in December 2022. Subsequently included publications were analyzed for content using VERBI’s qualitative data analysis software MAXQDA (version 2020). A list of the publications included can be found in [Supplementary Appendices A and B](#).

For the first review, articles were selected that used, presented, or described an instrument to measure pre-service or in-service science teachers’ enacted TPACK in lessons. In the second review, articles were selected that designated quality criteria for assessing the use of SGEVA in science education. Thus, any documents, including gray literature (such as conference proceedings) published in English between January 2010 and December 2022 were consulted for a comprehensive insight into

the current state of international research. Articles were excluded in the first review if the instruments were not designed to be applied to assess lesson planning, observation, and/or reflection. They were also ruled out if they focused on a specific subject domain outside the science subjects (e.g., English as a foreign language). In the second review, articles were excluded if they did not cover the creation of digital artifacts with moving images, if the artifacts were not produced by the students, or if they explicitly excluded the domain of science.

The search strategy for instruments measuring the enacted TPACK of PSSTs consisted of the following combination of keywords and Boolean operators: “TPACK” AND “practical” OR “enacted” OR “application” AND “science” OR “biology” AND “instrument” OR “rubric.” The search strategy utilized for documents that mentioned quality criteria related to SGEVA in science classes consisted of a combination of the following keywords and Boolean operators: “explainer video” OR “explanatory video” OR “animation” AND “student-generated” OR “student-created” OR “learner-generated” OR “learner-created” AND “science” OR “biology.” Search results were then exported into a spreadsheet program and screened according to the eligibility criteria and restrictions. Figure 1 presents the flow diagram for the first review while Figure 2 shows the flow diagram of the second review according to the Preferred Reporting Items for Systematic Reviews (PRISMA) 2020 statement (Page et al., 2021).

During the first review, the article titles and abstracts were screened according to the eligibility criteria, excluding 504 reports. Subsequently, 27 full texts were screened and another 6 records were excluded. No further cited publications were included via other methods because they did not meet the eligibility criteria. This left 21 records included in the first review.

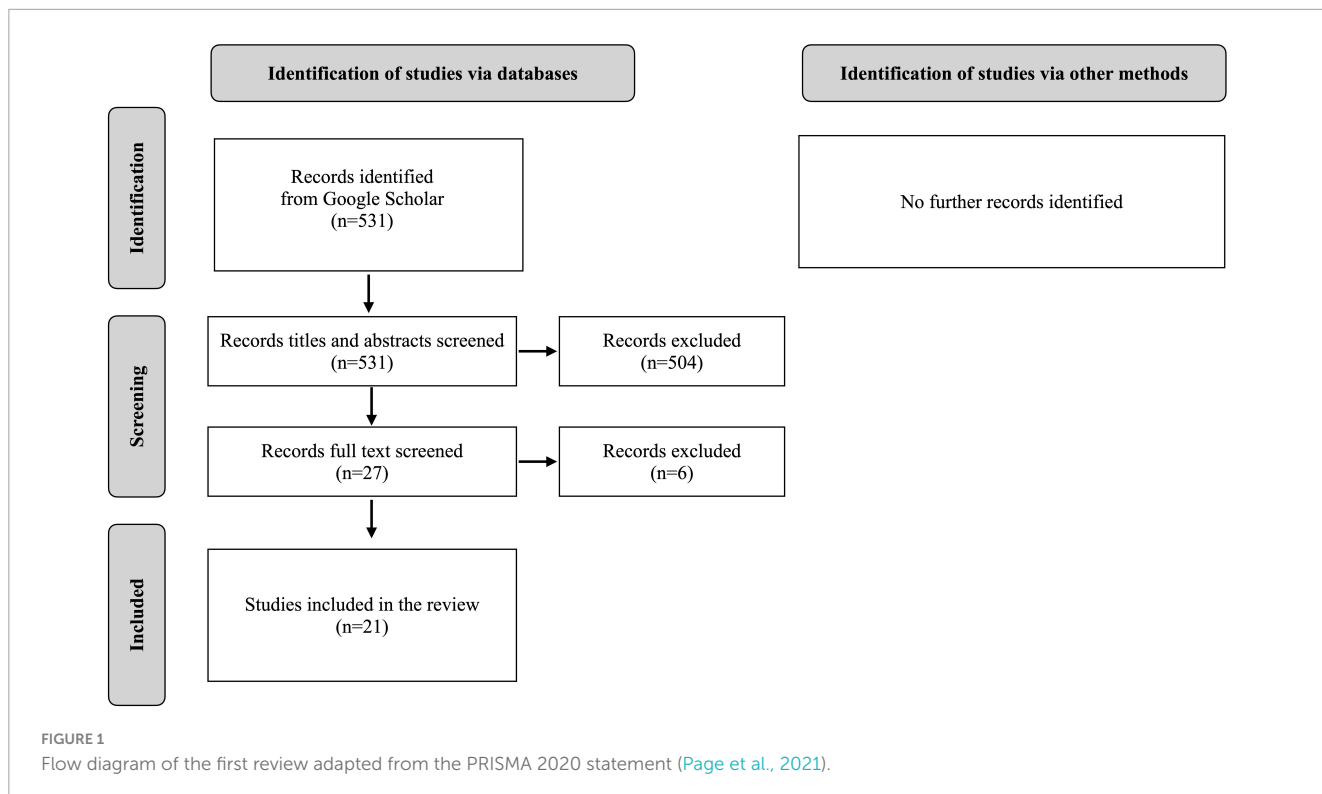
In the second review, the titles and abstracts of the documents were screened in terms of the eligibility criteria, excluding 952 reports. Next, the remaining 48 full texts were screened and another 19 records were excluded. Nine publications were then identified via citation screening of the remaining texts, leaving 38 records from the second review.

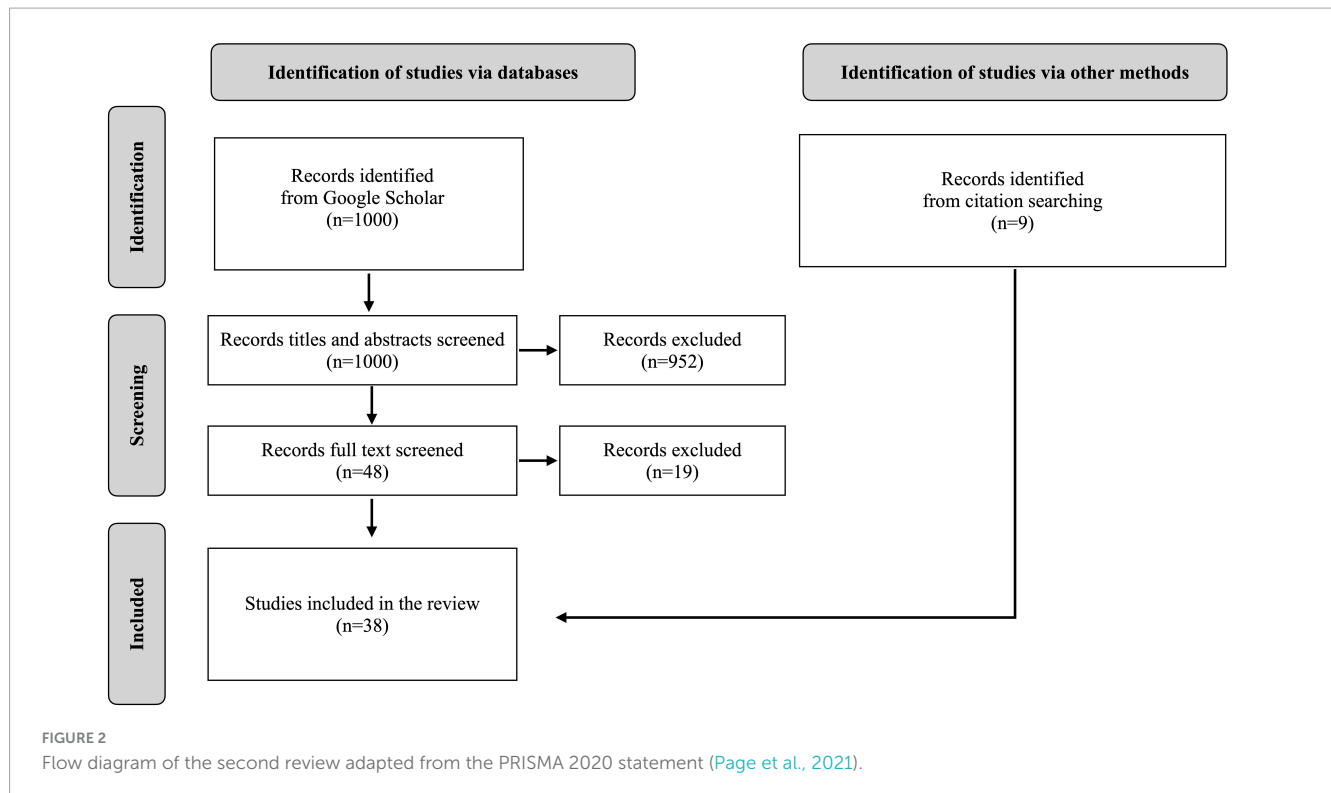
2.3. Development of the initial rubric

A draft version of the instrument was developed by comparing and subsequently structuring the results of the systematic review according to the theoretical basis. The results of the systematic review of literature on the enacted TPACK instruments and selected contextual factors were then used to guide the design of the rubric.

2.4. Evaluation: validity

The construct validity of the instrument was ensured by specifying the underlying concepts in terms of the literature reviewed (Döring, 2023). Additionally, the rubric’s content validity was empirically tested using judgments gathered online via www.sosicisurvey.de (version 3.2.50) from 11 experts in the field of digitally-enhanced science teaching. Each of the experts had at least a PhD in the field of science teaching and belongs to a scientific working group that explicitly deals with topics related to digitalization in science education. In the online survey, the rubric was systematically presented and differentiated into its individual levels (categories, subcategories, criteria, and indicators). The experts were asked for their feedback on the relevance of the categories. In parallel, they were able to





comment on each level (e.g., the individual indicators) within the categories.

2.5. Implementation: field trials

We tested the practicality of the instrument via a comprehensive case study in the classroom. To do this, we collected data sets from three PSSTs during their internship semester consisting of their planning outlines, the video recordings of their lesson implementation, and the recordings from their reflection interviews. All participants were in their master's studies, took part in a preparatory workshop, and stated that they had already conducted between 1 and 10 lessons with the support of DTM. Thus, the field trials correspond to the finalized procedure, which will represent the application context of the rubric. The data sets were evaluated by means of evaluative qualitative content analysis (Kuckartz and Rädiker, 2022) using the initial rubric. Thus, the deductively defined indicators were supplemented with inductively derived indicators.

2.6. Evaluation: reliability

Finally, the rubric was tested for reliability by comparing the analyses of two expert raters who edited 10% of the expected total data. As recommended by Al-Harhi et al. (2018), the second rater was first trained by the author in applying the rubric. Subsequently, the data sets (written lesson plans, videotaped lesson observations, and transcribed reflective interviews) of four cases were assessed by the two different raters, using the rubric. Both Cohen's Kappa coefficient and intraclass correlation coefficient

(ICC) were calculated using IBM SPSS (version 28.0). Cohen's Kappa was used to compare the ratings of the nominally scaled criteria before grading, whereas ICC was used to measure the inter-rater agreement of the intervally scaled levels (Döring, 2023).

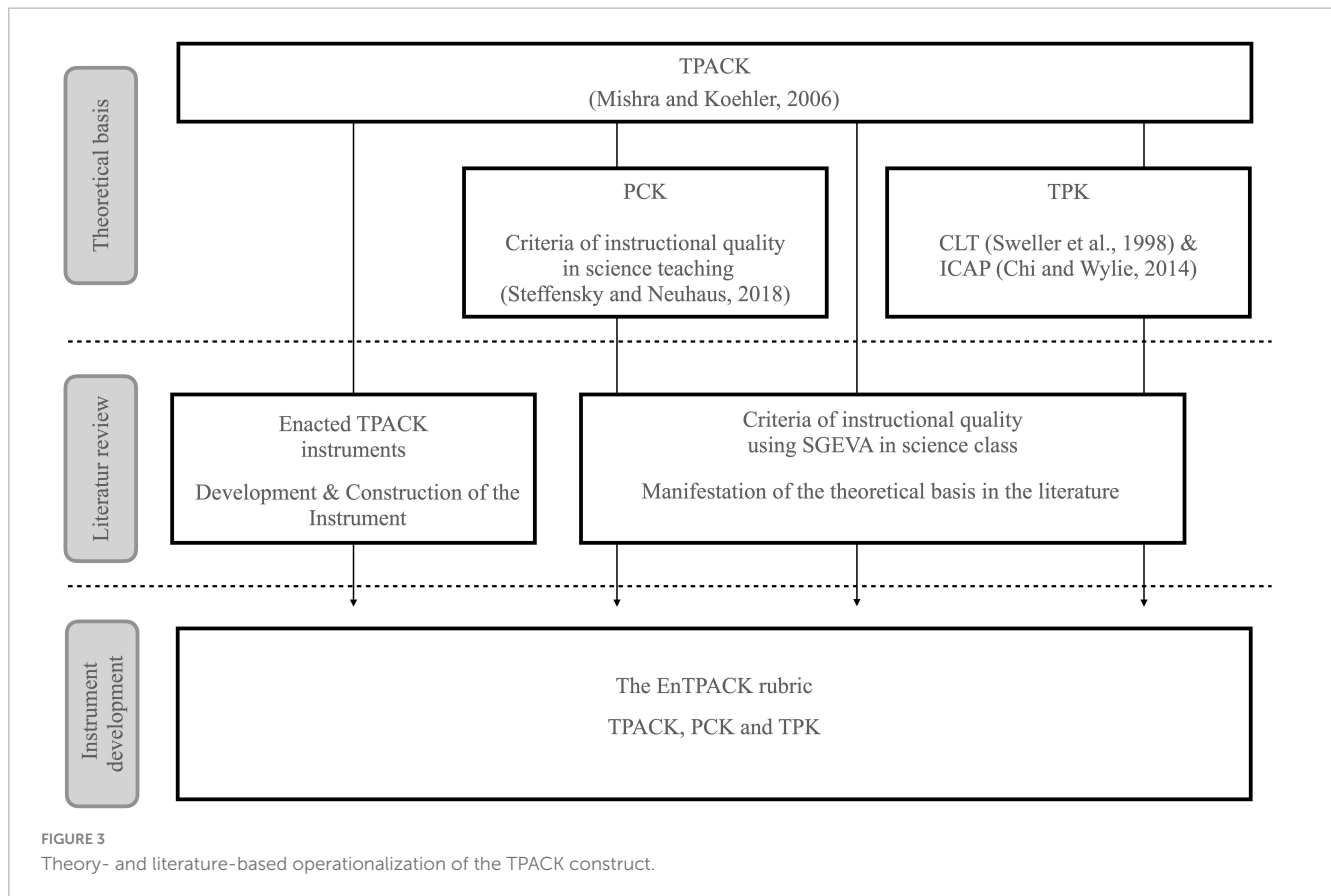
3. Results

In the following section, we present the theoretical basis of the rubric and the results of the two reviews, followed by the adjustments made during the validation and field trials. Finally, we discuss the reliability of the instrument.

3.1. Objective and theoretical basis of the rubric

The rubric was developed to measure the enacted TPACK of PSSTs who assigned SGEVAs in their internship semester at the levels of lesson planning, implementation, and reflection by using written planning artifacts (planning), video-recorded lesson observations (implementation), and reflection interviews (reflection). To address the situational and context-bound nature of the TPACK construct, the rubric is specified at the TK facet for explainer videos or animations, at the PK facet for SGEVA, and at the CK facet for the science domain.

Due to the application context of the instrument, the knowledge facets TPK, PCK, and TPACK were selected as central levels. The knowledge facets TK, PK, CK, and TCK were not further considered in the instrument structure for the following reasons: (TK) technological equipment is provided to the PSSTs



by the university if it is not available at the respective internship school. For this reason, as well as due to varying technological infrastructure at the schools, technological preparations cannot be adequately compared between subjects (PK). General pedagogical considerations such as classroom management allow only limited conclusions to be drawn in the context of the internship semester, since the PSSTs teach in their mentors' classes and usually adopt their pedagogical structures. In addition, overlapping PK aspects can be found in the PCK facet (CK). Prescribing a specific subject content was not possible in the internship semester for organizational reasons (TCK). The application of (digital) technologies in the science domain is not further relevant in the context of the present study.

To concretize the selected TPACK components, further theories were considered. For PCK, we used the criteria of instructional quality in science class developed by Steffensky and Neuhaus (2018). They define two subject-specific (basic) dimensions of deep structures that substantially influence the quality of science teaching. These dimensions are content structuring and cognitive activation, which are operationalized by exemplary indicators (e.g., feedback). For TPK, the Interactive-Constructive-Active-Passive model (ICAP; Chi and Wylie, 2014) and the Cognitive Load Theory (CLT; Sweller et al., 1998) were consulted. In the ICAP hypothesis, students' engagement within the learning process is divided into four levels. As students' engagement increases from passive to active, to constructive and lastly to interactive, their learning potentially increases accordingly (Chi and Wylie, 2014). The CLT, on the other hand, states that cognitive resources must be allocated wisely in learning processes in order to

focus sufficient capacity on engagement with the relevant learning contents (intrinsic and germane cognitive load; Sweller et al., 1998). Conversely, it is particularly important to reduce the cognitive load of irrelevant contents (extraneous cognitive load; Sweller et al., 1998). Both theories were selected since the planned activity turns students into DTM producers—which provides plenty of potential for interactive learning experiences (ICAP), but also presents a risk of cognitive overload (CLT). Figure 3 depicts the operationalization of the TPACK construct conducted in the present study.

3.2. Literature review on enacted TPACK instruments

Overall, twelve of the twenty-one articles (57.14%) had been published in journals, seven (33.33%) in conference proceedings, and two (9.52%) as book chapters. Figure 4 illustrates the increase in publications regarding the measurement of TPACK based on lesson planning, implementation, and reflection in the course of the past decade.

Seven publications assessed enacted TPACK without a specific subject focus. Six articles focused on science in general, four on biology, two on chemistry, one on the combination of biology and physics, and another one on geology. Sixteen different instruments for measuring enacted TPACK using various methods and covering different data sources were used or described. We gained insight into 14 of these 16 instruments, most of which were included in the corresponding publications, others were publicly available

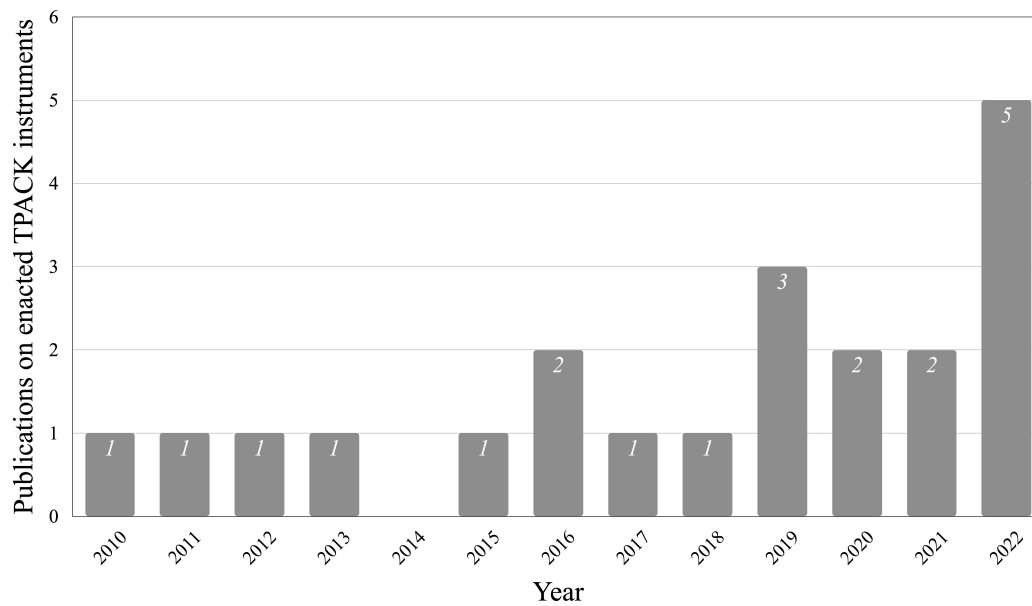


FIGURE 4

Annual number of publications on enacted TPACK instruments, January 2010 to December 2022 (inclusive).

online (Harris et al., 2010, 2012; Hofer et al., 2011), or were provided on request by their authors (Canbazoglu Bilici et al., 2016). The instruments varied in terms of the data sets to which they were applied, their structure and handling, the process of their development, their degree of specificity in terms of the TPACK facets or contextual factors, and their demonstrable validity and reliability. Different methods and instruments were often combined but such combinations lay outside the scope of this investigation. Table 1 provides a rough summary of the review, which will be concretized within the individual subchapters.

3.2.1. Data sets or combinations used

To assess enacted TPACK, instruments used to analyze lesson planning, lesson implementation, and lesson reflection were considered. For lesson planning, written planning documents or other texts (e.g., Harris et al., 2010; Al-Harathi et al., 2018; Pareto and Willermark, 2019; Hayati et al., 2022; Unal Coban et al., 2022) and audio-recorded planning interviews (e.g., Harris et al., 2012; Pareto and Willermark, 2019) were used. In addition, some teacher-generated materials (e.g., presentations, handouts, or digital artifacts) were included in the analysis (e.g., Al-Harathi et al., 2018; Pareto and Willermark, 2019; Unal Coban et al., 2022).

To analyze lesson implementation, direct observational methods were used by the researchers (e.g., Kafyulilo et al., 2013; Pareto and Willermark, 2019; Paidi et al., 2020; Kawai and Nyamupangedengu, 2021), who often took field notes or used log books to supplement their observations. In contrast, some approaches used videotaped lesson observations (e.g., Hofer et al., 2011; Yeh et al., 2015; Canbazoglu Bilici et al., 2016; Aktaş and Özmen, 2020, 2022). To support the assessment of lesson implementation, materials produced for and by the students or personal conversations were occasionally utilized (Pareto and Willermark, 2019). Lesson reflections were collected via

questionnaires (Kafyulilo et al., 2013) or observations of reflection sessions and interviews (Pareto and Willermark, 2019).

Most instruments focused on measuring enacted TPACK by exclusively analyzing lesson planning (e.g., Harris et al., 2010, 2012; Janssen and Lazonder, 2016; Özdilek and Robeck, 2019; Dewi et al., 2022) or lesson implementation (e.g., Hofer et al., 2011; Kafyulilo et al., 2013; Paidi et al., 2020; Kawai and Nyamupangedengu, 2021). However, combined analysis using one instrument was relatively rare, with only three approaches reporting on the combined measurement of lesson planning and lesson implementation (Yeh et al., 2015; Pareto and Willermark, 2019; Aktaş and Özmen, 2020, 2022). However, only Yeh et al. (2015) differentiated indicators for planning from those for implementation—and did so only partly.

Some researchers adapted their instruments to be used for different data sets (e.g., Harris et al., 2010, 2012; Hofer et al., 2011; Canbazoglu Bilici et al., 2016). Canbazoglu Bilici et al. (2016) developed two instruments: the TPACK-based lesson plan assessment instrument (TPACK-LpAI) to analyze written lesson plans and the TPACK Observation Protocol (TPACK-OP) to analyze videotaped lesson observations. These two instruments use the same key indicators and differ only slightly, as the TPACK-OP incorporates contextual factors such as classroom demographics and setting in greater depth. Harris et al. (2010) developed a broadly designed instrument for written lesson plans, which they subsequently adapted for use in lesson observations (Hofer et al., 2011) and for in-service teacher lesson planning interviews (Harris et al., 2012). Again, these instruments differ in only a few respects. Consequently, two criteria that explicitly refer to technology logistics and observed instructional use were added for lesson observations. Furthermore, contextual factors and aspects that were not expressed during lessons were collected in an additional field notes column. No changes were made when using the instrument in lesson planning interviews. The research team flexibly used the instrument to provide valid and

TABLE 1 Results of the first review (enacted TPACK instruments).

References	Application to data sets			Structural orientation to the TPACK construct			Specification of the instrument (following Njiku et al., 2020)			Level of contextualization (following Brianza et al., 2022)			Tested for validity and reliability	
	Planning	Implementation	Reflection	All Facets	Selected Facets	Overall construct	TK	PK	CK	Immediate	Proximal	Distal	Validity	Reliability
Aktaş and Özmen, 2020	×	×	-	-	-	×	-	-	×	×	-	×	×	×
Aktaş and Özmen, 2022	×	×	-	-	-	×	-	-	×	×	-	×	×	×
Al-Harathi et al., 2018	×	×	-	-	-	×	×	-	-	×	-	-	×	×
Antony and Paidi, 2019	-	×	-	-	×	-	-	-	(?)	-	-	×	×	×
Canbazoğlu Bilici et al., 2016	×	×	-	-	-	×	-	-	×	-	-	×	×	×
Dewi et al., 2022	×	-	-	×	-	-	-	-	-	-	-	-	-	-
Harris et al., 2010	×	-	-	-	-	×	-	-	-	-	-	×	×	×
Harris et al., 2012	×	-	-	-	-	×	-	-	-	-	-	×	×	×
Hayati et al., 2022	×	-	-	×	-	-	-	-	(?)	-	-	-	-	-
Hofer et al., 2011	-	×	-	-	-	×	-	-	-	-	-	×	×	×
Janssen and Lazonder, 2016	×	-	-	×	-	-	-	×	×	×	-	-	-	×
Kafyulilo et al., 2013	-	×	-	-	×	-	-	-	×	-	-	-	-	×
Kawai and Nyamupangedengu, 2021	-	×	×	-	×	-	-	-	-	-	-	-	×	×
Mourlam et al., 2021	×	-	-	-	-	×	-	-	-	-	-	×	×	×
Mourlam and Bleecker, 2017	×	-	-	-	-	×	-	-	-	-	-	×	×	×
Özdilek and Robeck, 2019	×	-	-	-	×	-	-	×	×	×	-	-	-	-

(Continued)

TABLE 1 (Continued)

References	Application to data sets			Structural orientation to the TPACK construct			Specification of the instrument (following Njiku et al., 2020)			Level of contextualization (following Brianza et al., 2022)			Tested for validity and reliability	
	Planning	Implementation	Reflection	All Facets	Selected Facets	Overall construct	TK	PK	CK	Immediate	Proximal	Distal	Validity	Reliability
Paidi et al., 2020	-	x	-	-	x	-	-	-	(?)	-	-	-	x	x
Pareto and Willermark, 2019	x	x	-	-	x	-	-	x	-	x	-	-	-	-
Unal Çoban et al., 2022	x	-	-	-	-	x	-	x	-	-	-	-	x	x
von Kotzebue, 2022	x	-	-	-	-	x	-	x	-	-	-	-	-	x
Yeh et al., 2015	x	x	-	-	-	x	-	x	-	-	-	-	-	x

"x" = fulfilled "x" = not fulfilled "(?)" = no reliable information available.

reliable measurements of three different data sets. The authors concluded that the instrument was suitable for the analysis of oral interviews about technology-enhanced lessons, which would also include its use for reflecting orally on lessons (Harris et al., 2012). Similarly, Canbazoğlu Bilici et al. (2016) asserted that the TPACK-OP rubric could be used to assess lesson plan reflections.

Nevertheless, the review identified no instrument that was explicitly designed or used for the combined analysis of lesson planning, implementation, and reflection. In general, lesson reflections were evaluated separately to obtain more in-depth information (Kafyulilo et al., 2013; Pareto and Willermark, 2019).

3.2.2. Instrument structure

The included instruments can be described as rubrics, as they are composed of different assessment categories divided into rating scales for assessing quality (Popham, 1997). Among the instruments, the number of categories ranged from four in Harris et al. (2010) to twelve in Aktaş and Özmen (2020, 2022) who further differentiated the categories into several indicators. Most instruments were not structurally oriented to particular facets of TPACK but defined categories that allow overall conclusions to be drawn about TPACK as a construct (e.g., Yeh et al., 2015; Al-Harathi et al., 2018; Unal Çoban et al., 2022). Only a few rubrics used all TPACK components for categorization (e.g., Hayati et al., 2022), while others exclude individual facets (Kafyulilo et al., 2013), and focused either on the integrated components (TPK, TCK, and PCK; Janssen and Lazonder, 2016; Paidi et al., 2020) or the basic ones (TK, PK, and CK; Pareto and Willermark, 2019; Kawai and Nyamupangedengu, 2021).

Every instrument used rating scales containing three (e.g., Kafyulilo et al., 2013) to five (e.g., Canbazoğlu Bilici et al., 2016) levels, mostly scoring each category separately and in some cases even allowing for variable scoring within the categories (von Kotzebue, 2022). Three instruments did not score every category but calculated an overall score by mapping interactions between coded TPACK categories (Pareto and Willermark, 2019; Kawai and Nyamupangedengu, 2021) or using a calculation formula (Paidi et al., 2020).

Accordingly, the scoring methods varied among the different approaches and were (in part) highly dependent on inference. For example, some instruments used a Likert scale (e.g., Kafyulilo et al., 2013; Canbazoğlu Bilici et al., 2016; von Kotzebue, 2022), while others provided concrete descriptors for each category level (e.g., Harris et al., 2010, 2012; Hofer et al., 2011; Yeh et al., 2015; Al-Harathi et al., 2018; Özdilek and Robeck, 2019).

3.2.3. Development of the instruments

The development process for two of the instruments was not described (Özdilek and Robeck, 2019; Dewi et al., 2022). However, most researchers adapted the items or indicators of previously established technology integration assessment instruments (Harris et al., 2010), TPACK-instruments (e.g., Kafyulilo et al., 2013; Aktaş and Özmen, 2020, 2022; Unal Çoban et al., 2022), the integrative PCK model (Canbazoğlu Bilici et al., 2016), or other theoretical models (von Kotzebue, 2022). Literature reviews were used to develop a few more specialized instruments, although these were not systematically documented

(e.g., Al-Harathi et al., 2018; Unal Çoban et al., 2022). Other authors built on self-developed operationalizations of enacted TPACK (Yeh et al., 2015; Pareto and Willermark, 2019). For example, Yeh et al. (2015) conducted a Delphi study to operationalize the enacted TPACK of in-service teachers with DTM experience via interviews, subsequently testing and revising the instrument built on this operationalization. Field trials were also used by other researchers to test their instruments in practice (e.g., Al-Harathi et al., 2018; Antony and Paidi, 2019).

3.2.4. Specification of the TPACK components

The instruments vary widely in terms of specificity at the levels of technology, pedagogy, and content. Some were designed broadly to fit a wide range of approaches without any specification (e.g., Harris et al., 2010, 2012; Hofer et al., 2011; Dewi et al., 2022). One instrument was designed expressly for a specific use of DTM (cloud-based learning activities) specializing only at the TK level (Al-Harathi et al., 2018). Most were specifically designed for the domain of science (CK) in general (e.g., Kafyulilo et al., 2013; Yeh et al., 2015; Canbazoglu Bilici et al., 2016; Aktaş and Özmen, 2020, 2022), since this was one of the eligibility criteria used to select them for review. This subject-specific focus was usually manifested through generic PCK categories of high-quality science teaching, which are either evaluated for the use of DTM (Canbazoglu Bilici et al., 2016) or supplemented with generic TPK categories (von Kotzebue, 2022). However, there are also instruments focusing on both the subject domain (CK) and a specific instructional approach (PK), such as inquiry learning in biology (Janssen and Lazonder, 2016), case-based learning in chemistry (Özdilek and Robeck, 2019) or argumentation in the geology classroom (Unal Çoban et al., 2022).

Although some studies focused on particular session content (e.g., von Kotzebue's (2022) investigation of a lesson about honeybees) or together with a specific DTM use (e.g., glucose-insulin regulation with System Dynamics modeling; Janssen and Lazonder, 2016), these highly specific aspects were not reflected in the corresponding research instruments.

3.2.5. Contextualization

Brianza et al. (2022) conceptualize educational context as an interdependent network of contextualized technology-, pedagogy-, and content-related factors differentiated on three different levels of influence to classroom practice and the teachers' realm of influence (immediate, proximal, and distal). On the class level, the teacher can immediately influence contextual conditions, whereas his realm of influence is proximal on the school or institution level and distal on national or global level.

In the reviewed instruments, these contextual factors were integrated in different ways. Either the instruments were designed with these factors in mind or items/indicators were integrated that specifically address contextual factors and thus include them in the measurement process. The most frequently identified contextual factors included in the reviewed instruments represent distal pedagogy- or content-related factors. Two approaches mentioned the incorporation of national teacher qualification standards in the design of their instruments (Antony and Paidi, 2019; Aktaş and Özmen, 2020, 2022). Several other instruments assessed the

compatibility of the lesson with national curriculum standards (e.g., Harris et al., 2010, 2012; Hofer et al., 2011; Canbazoglu Bilici et al., 2016). Immediate contextualized pedagogy-related factors, such as teacher or student dispositions, were mostly integrated among generic PK or PCK indicators, that included, for example, the activation of prior knowledge (Janssen and Lazonder, 2016), the adaptation to different learning preferences (Al-Harathi et al., 2018), the consideration of students' needs (Unal Çoban et al., 2022), or the arousal of attention and the activation of the students (Aktaş and Özmen, 2022). Another example for the inclusion of immediate pedagogy- and technology-related contextual factors via items was the instrument of Canbazoglu Bilici et al. (2016), who added the class grade and size, the arrangement of the room, and the availability of DTM. Particularly rare was the inclusion of contextual factors on the proximal school or institutional level, which was identified solely by focusing the instrument on a particular type of school (e.g., primary schools; Pareto and Willermark, 2019).

3.2.6. Validation and reliability

In general, the instruments were tested for reliability more often than for validity. Four articles commented on neither the validity nor the reliability of their instruments (Pareto and Willermark, 2019; Özdilek and Robeck, 2019; Dewi et al., 2022; Hayati et al., 2022) although Pareto and Willermark (2019) reported comparing ratings from two different coders.

To measure reliability, all the remaining instruments used either the percent agreement procedure (e.g., Yeh et al., 2015; Janssen and Lazonder, 2016), the intraclass correlation coefficient (e.g., Harris et al., 2010, 2012; Hofer et al., 2011; Canbazoglu Bilici et al., 2016; Antony and Paidi, 2019), Cohen's Kappa (e.g., Kafyulilo et al., 2013; Janssen and Lazonder, 2016; Al-Harathi et al., 2018; von Kotzebue, 2022), or Kendall's W (e.g., Aktaş and Özmen, 2020, 2022; Unal Çoban et al., 2022).

To establish face or construct validity, the researchers primarily sought expert judgments (e.g., Harris et al., 2010, 2012; Hofer et al., 2011; Canbazoglu Bilici et al., 2016; Al-Harathi et al., 2018) with two articles reviewing the literature as a supplementary validation of content (Al-Harathi et al., 2018; Kawai and Nyamupangedengu, 2021). One study measured empirical validity using biserial correlation value (Antony and Paidi, 2019). In addition, test-retest reliability was calculated using Pearson's r (Canbazoglu Bilici et al., 2016) and internal consistency using Cronbach's Alpha (e.g., Kafyulilo et al., 2013; Canbazoglu Bilici et al., 2016; Al-Harathi et al., 2018; Aktaş and Özmen, 2020, 2022).

Besides the review of enacted TPACK instruments, a second review was carried out, which focused on quality criteria for the use of SGEVA in the classroom. In the following chapter, the results of this second review are presented.

3.3. Literature review on student-generated explainer videos and animations in science classes

Among the 38 articles, 29 were published in journals (76.3%), 3 in conference proceedings (7.9%), 4 in book chapters (10.5%), and 2 as dissertations (5.2%). The three journals that featured most often

in the sample were *Chemistry Education Research and Practice*, with three articles, the *Journal of Biological Education*, and the *Journal of Science Education and Technology*, with two articles apiece.

The included articles varied in regard to the educational levels examined, the subject areas explicitly focused on, and the products created by the students. It should be noted that some articles mention multiple educational levels, which may result in overlaps. Among the 38 articles reviewed, 18 focused on tertiary education, with 10 of those specifically targeting pre-service teachers; the articles focusing on secondary and primary education numbered 15 and 7, respectively while 5 articles did not explicitly mention any educational level.

Overall, 18 articles (47.3%) focused on science education while 6 articles each (15.8%) covered biology or chemistry education or did not explicitly mention any subject. Another two articles (5.3%) identified STEM education as their focus. In total, 19 articles (50%) examined the creation of animations by the students and 10 of these dealt with a special slowed-down form called “slowmation” (Hoban and Nielsen, 2010). 14 articles (36.8%) dealt with student-created videos and 5 articles (13.2%) involved the creation of various digital media products, explicitly including videos and/or animations. The following section clusters the criteria identified in the review according to the selected theoretical basis.

3.3.1. TPACK

Consistent with the TPACK framework (Mishra and Koehler, 2006), a need for alignment of the DTM use with both the selected subject content and the learning objectives is highlighted in the literature. First, not every use of DTM is suitable for each method and for teaching all topics. Accordingly, the appropriateness of the selected subject content for SGEVA needs to be carefully evaluated (Keast and Cooper, 2011). To create a short explainer video, the subject content needs to be limited in scope, freestanding, and divisible into visualizable steps (e.g., Keast and Cooper, 2011; Farrokhnia et al., 2020; Orraryd, 2021). In particular, dynamic content that involves motion or spatio-temporal change benefits from being represented in moving images (e.g., Wu and Puntambekar, 2012; Hoban and Nielsen, 2014; Akaygun, 2016; Yaseen, 2016, 2018). In addition, complex and abstract processes or phenomena that are invisible to the human eye benefit from visualization (e.g., Kidman et al., 2012; Akaygun, 2016; Orraryd, 2021). This does not mean that SGEVA do not suit for example static subject content, but with regard to subject content with the described characteristics, the potential of SGEVA is maximized.

Second, DTM use must be aligned with subject-specific learning objectives (Reyna and Meier, 2018a). SGEVAs are therefore particularly useful for consolidating previously exposed subject content (e.g., Hoban and Nielsen, 2010; Keast and Cooper, 2011; Hoban, 2020). In this regard, the focus must be on the scientific process or concept while the DTM should serve as a subject-learning tool (e.g., Hoppe et al., 2016; Seibert et al., 2019; Nielsen et al., 2020; Orraryd, 2021). Accordingly, the process (not the product) of video creation should be foregrounded (e.g., Keast and Cooper, 2011; Potter et al., 2021).

3.3.2. PCK–content structuring

Instructional quality in science lessons can be achieved by considering content structuring (Steffensky and Neuhaus, 2018). In

the reviewed articles, this can be accomplished (1) when content is taught in a structured and adequately supported way, (2) when technical language is considered in the task, and (3) when the teacher communicates the learning objectives transparently. To consider the first point, supportive teaching should provide students with content-related (e.g., Hoppe et al., 2016; Orraryd, 2021), criteria-based (e.g., Kamp and Deaton, 2013; Andersen and Munksby, 2018; He and Huang, 2020) and appreciative (e.g., Yaseen and Aubusson, 2018; Orraryd, 2021) feedback. Appropriate structuring includes emphasizing key aspects of the subject content over minor issues (e.g., White et al., 2020; Orraryd, 2021) and coherently linking the task to students’ prior knowledge (e.g., Kidman et al., 2012; Farrokhnia et al., 2020; Nielsen et al., 2020; Orraryd, 2021). Second, the use of technically correct language (Seibert et al., 2019) when supplementing moving images with verbal explanation (e.g., Karakoyun and Yapıcı, 2018; Farrokhnia et al., 2020) is considered another important aspect of SVEGA. Third, teachers should clearly communicate learning objectives and define expectations for students (e.g., Reyna and Meier, 2018b; He and Huang, 2020; Nielsen et al., 2020). Accordingly, students should be made aware of the purpose and the opportunities of SGEVA in terms of learning objectives (e.g., Lawrie and Bartle, 2013; Andersen and Munksby, 2018; Epps et al., 2021).

3.3.3. PCK–cognitive activation

Cognitive activation is another important characteristic of instructional quality in science teaching (Steffensky and Neuhaus, 2018). According to the reviewed articles this is expressed in (1) task openness, (2) authenticity, and (3) emphasis on content relevance. In terms of *task openness*, encouraging student creativity as well as co-construction and allowing a certain degree of artistic freedom are essential while also requiring conceptual fidelity (Orraryd, 2021). Students should not simply be given ready-made materials and a pre-prepared storyline (Robeck and Sharma, 2011; Mills et al., 2019), but be encouraged to make their own decisions about design, format, and storytelling (e.g., Robeck and Sharma, 2011; Farrokhnia et al., 2020; Epps et al., 2021; Potter et al., 2021) while supporting them adequately in problem solving (Palmgren-Neuvonen and Korkeamäki, 2015) and content-related issues (Mills et al., 2019; Orraryd, 2021). Providing open-ended fictive tasks (Palmgren-Neuvonen et al., 2017) or allowing students to conduct their own investigations (Hoppe et al., 2016) are mentioned as examples in this context.

To increase *task authenticity*, teachers should provide an objective for the SGEVA (e.g., Willmott, 2014; Pirhonen and Rasi, 2016; Blacer-Bacolod, 2022). For instance, explainer videos might be uploaded to an open or closed platform (e.g., Hoppe et al., 2016; Reyna and Meier, 2018a; Baclay, 2020; Gallardo-Williams et al., 2020). In addition, a realistic audience for the video should be provided (Ribosa and Duran, 2022). The teacher needs to make the students aware of their audience and guide them in this regard (e.g., Reyna and Meier, 2018a; Gallardo-Williams et al., 2020; Hoban, 2020). Finally, the *real-life relevance of the subject content* needs to be addressed (Seibert et al., 2019). A well-designed task provides a real-life problem or question (Hoppe et al., 2016) concerning a topic of relevance to students’ lives (e.g., Farrokhnia et al., 2020; Epps et al., 2021).

3.3.4. TPK–cognitive load

Most of the articles recommended that practitioners provide various forms of scaffolding through different agents such as teachers, learning tools, or instructional materials (e.g., Wu and Puntambekar, 2012) and align in this respect with the CLT (Sweller et al., 1998). Essentially, three areas of extrinsic cognitive load are addressed: (1) the use of technology, (2) the planning and design of the video, and (3) the task organization and management.

Regarding the *use of technology*, teachers must take into account the students' media-related skills (e.g., Reyna and Meier, 2018b; Farrokhnia et al., 2020; Epps et al., 2021; Blacer-Bacolod, 2022). The video production software should be user-friendly while providing sufficient editing options (e.g., Belski and Belski, 2014; Karakoyun and Yapıcı, 2018). Teachers need to introduce mandatory functions (e.g., Hoban and Nielsen, 2014; Seibert et al., 2019; Epps et al., 2021) and support students' use of the technology (e.g., Yaseen, 2016; Reyna and Meier, 2018a,b).

Storyboards are widely recommended to support the *video design process* (e.g., Robeck and Sharma, 2011; Pirhonen and Rasi, 2016; Palmgren-Neuvonen et al., 2017; Orraryd, 2021). Also recommended are introducing video production techniques (e.g., Seibert et al., 2019; Hoban, 2020), giving demonstrations with exemplary videos or animations (e.g., Kidman et al., 2012; Kamp and Deaton, 2013; Palmgren-Neuvonen and Korkeamäki, 2014; Hoban, 2020), and providing guidance on video length and structure (Gallardo-Williams et al., 2020). While some authors also recommend training students in digital media principles (Reyna and Meier, 2018b), others emphasize embracing the imperfections of the videos (e.g., Yaseen, 2018; Orraryd, 2021).

Since *organization and task management* increase with group size, several authors propose forming small groups consisting of three to four members (e.g., Yaseen, 2016, 2018; Jacobs and Cripps Clark, 2018; Seibert et al., 2019). To create a healthy and productive environment, group composition should be considered. In this respect, it is recommended to consider students' personal relationships when forming groups (Kidman et al., 2012) and to ensure that each group contains a variability of skills (e.g., Robeck and Sharma, 2011; Palmgren-Neuvonen and Korkeamäki, 2014). In addition, it is advantageous to clearly define roles and subtasks (e.g., Keast and Cooper, 2011; Karakoyun and Yapıcı, 2018; Reyna and Meier, 2018a; Epps et al., 2021) and the teacher must mediate intra-group conflicts when required (e.g., Palmgren-Neuvonen and Korkeamäki, 2014, 2015; Palmgren-Neuvonen et al., 2017; Karakoyun and Yapıcı, 2018; Reyna and Meier, 2018a,b).

3.3.5. TPK–interactive learning

Consistent with the ICAP model (Chi and Wylie, 2014), SGEVAs should enable co-construction, collaboration, and communication among class participants (e.g., Jacobs and Cripps Clark, 2018; Özdilek and Ugur, 2019; Baclay, 2020; Blacer-Bacolod, 2022). Students should solve problems, brainstorm, or plan collaboratively via teacher-student or peer exchange (e.g., Robeck and Sharma, 2011; Kamp and Deaton, 2013; Mills et al., 2019; Seibert et al., 2019) leading to content-related discussions. Such processes can be aided by the teacher as they occur (e.g., Hoban and Nielsen, 2014; Palmgren-Neuvonen and Korkeamäki, 2014; Yaseen, 2018; Farrokhnia et al., 2020) by promoting cognitive conflicts, asking key questions (Yaseen, 2016; Yaseen and Aubusson, 2018),

requesting students to support their assertions (Yaseen, 2016, 2018; Palmgren-Neuvonen et al., 2017; Orraryd, 2021), and summarizing or highlighting contributions in the discussion (Yaseen, 2016; Yaseen and Aubusson, 2018). When implementing the videos, the students should be cognitively activated; for example, by being encouraged to represent the subject content multimodally (Lawrie and Bartle, 2013; Reyna and Meier, 2018a).

3.3.6. TPK–time allocation

The need to allocate sufficient time for video preparation and creation is often mentioned (e.g., Özdilek and Ugur, 2019; Epps et al., 2021; Blacer-Bacolod, 2022). This time varies from two (Seibert et al., 2019) to three hours (Hoban and Nielsen, 2014).

3.4. Development of the initial rubric

Following the SGEVA review, the elaborated results were compared with the selected theories and organized according to the structure inherent in the theories. The only exception to this process concerned "time allocation," which was separately assigned to TPK since it is causally attributed to planning and creating videos or animations. Our rubric was based on the results of the SGEVA review and included only those instructional criteria that appeared in the literature, thereby ensuring the rubric would be context-specific.

To operationalize the theory-driven subcategories, we used the cluster of criteria presented in section 3.3. The structure of the rubric divided subcategories such as cognitive load into clear criteria based on the results of the SGEVA review. Each criterion was further divided into observable indicators to ensure consistency. In the coding manual, we formulated explicit examples for each indicator to facilitate the process of rating the three different data sets. Based on how frequently they appeared in the review, we assigned a weight ranging from one to three to each indicator. We then created a four-level classification of the criteria that assigned grades based on a points system tailored to the total weighting of the associated indicators. Thus, the indicators were first observed before categorizing criteria into levels ranging from 0 (not fulfilled) to 3 (adequately fulfilled), based on the presence and weight of the observed indicators. The total scores could be calculated based on these criterial values.

Our focus on SGEVA in science classes meant we already had a contextualized focus regarding the TPACK facets. The additional contextual factors primarily represented the framework conditions of the project, including the situation of the PSSTs internship semesters and the conditions of their classroom observations. During the internship semester, the PSSTs visit their mentors' classes and thus do not influence classroom practices etc. Moreover, the classroom observation conditions restricted the time available to the procedure. The lesson observations thus only covered one double lesson (90 min), which made it impossible to achieve the recommended requirement of at least 2–3 h of planning and video creation within the project framework. The indicators were therefore adjusted according to this time limit; bearing in mind the possibility of continuing the activities in subsequent lessons, the specified timespan was reduced to 40 min.

3.5. Validity

Based on the expert feedback, the initial rubric was adapted. As a major change, redundancies were removed by merging subcategories with their associated criteria if they contained only one criterion. As a result, four of the seven subcategories (DTM aligned with content, DTM aligned with learning objective, Interactive learning, and Time allocation) were not subdivided into different criteria.

3.6. Field trials

The case study findings were subsequently used to trial the rubric in practice. We reformulated the indicators, adjusted the weightings, inductively added more indicators, and revised examples of the individual indicators in the coding manual. To increase objectivity, the reformulations served to operationalize the indicators as clearly as possible considering the actual data sets. This clarified any formulations that were misleading or open to interpretation. An indicator was reweighted if its frequency of use in the SGEVA review was associated with an excessively high or low influence on the grading of the criterion since this would bias the assessment.

Indicators could be included if they were not explicitly mentioned in the reviews but had a significant impact on teaching quality in practice. An example is the distinction between essential and non-essential technical terms in the subcategory of content structuring or the appropriateness of time allocation to the learning objectives. The rubric comprises the three categories of TPACK, PCK, and TPK (see [Table 2](#)), broken down into a total of seven subcategories. While four of these were sufficiently differentiated, three were further divided into three to four criteria each, resulting in a total of fourteen criteria. By assigning approximately five observable indicators, the classification of these criteria was based on the data sets (written lesson plans, video-recorded lesson observations, and transcribed lesson reflection interviews).

3.7. Reliability

The reliability of the instrument was assessed via two measures. The Cohen's Kappa of the instrument across all three datasets was 0.704, indicating a substantial inter-rater agreement ([Landis and Koch, 1977](#)). To measure inter-rater-reliability that included the grading process, the criteria levels were calculated based on the weighting of each indicator and measured using ICC. The resulting value of 0.873 indicated good overall reliability ([Koo and Li, 2016](#), see [Table 3](#)).

4. Discussion and conclusion

The overall objective of this study was to develop and validate an instrument for assessing PSSTs enacted TPACK based on planning, implementation, and reflection on a lesson using SGEVA. This objective raised questions about the development and construction of existing instruments assessing the enacted TPACK

of science teachers and about the criteria for judging successful teaching with SGEVA in science classes.

4.1. Review of enacted TPACK instruments

We screened 21 articles describing instruments for measuring enacted TPACK, identifying 16 different rubrics. Taking the bibliographic data into account, we observed an increase in the development of instruments for enacted TPACK assessment, in line with [Wang et al. \(2018\)](#). Nevertheless, there was a clear shortage of rubrics designed to measure and compare enacted TPACK using multiple data sets. Although several authors indicated that their rubrics could be applied to lesson reflections, none had used them in this way. However, reducing the enactment of TPACK to only one of the three phases of teaching (planning, implementation, and reflection) excludes aspects potentially relevant to professional development ([Canbazoglu Bilici et al., 2016](#); [Willermark, 2018](#)). This is especially the case for the reflection of the lesson, as this is considered central to the development of the teachers' professional competencies.

Most of the identified rubrics assessed TPACK on a rather broad level with a lack of clear descriptions and were largely undifferentiated. This requires graders to infer the precise meaning of criteria and leads to a vague operationalization of the fuzzy construct ([Brantley-Dias and Ertmer, 2014](#); [Willermark, 2018](#)). To situate the rubrics on specific conditions, supplementary literature was consulted on focused DTM (TK), instructional approaches (PK), or specific subject domains (CK). However, for none of the situated rubrics identified in this review, the review process was systematically documented. Additionally, the resulting categories were rarely mapped to the TPACK components. Given the lack of conceptual fidelity ([Wang et al., 2018](#)), it seemed reasonable to perform a situational mapping of the framework considering focused situational conditions.

The rubric presented in this paper addresses these limitations via its explicit design for application to planning, implementation, and reflection on science classes using student-generated explainer videos or animations. It operationalizes the TPACK construct via clearly described indicators that are based on existing theory and on empirical studies.

4.2. Review of SGEVA

To understand the focused use of DTM in this classroom approach, we reviewed 38 articles that provided criteria for assessing teaching with SGEVA in science classes. This process allowed us to gather a set of relevant instructional criteria based on the current literature.

The identified criteria were consistent with our theoretical basis. In line with the ICAP model ([Chi and Wylie, 2014](#)) the reviewed articles advocate for interactive as well as co-constructive learning whereas scaffolding on different areas is recommended consistent with the CLT ([Sweller et al., 1998](#)). Aligning DTM use with learning objectives and course content is described as a crucial factor in line with the TPACK

TABLE 2 The structure of the EnTPACK rubric.

Categories (3)	Sub-categories (7)	Criteria (14)	Indicators (68)	Exemplary indicator
TPACK	DTM aligned with content		6 indicators	(TPACK-f.d.) The selected subject content shows spatial and/or temporal changes.
	DTM aligned with learning objective		7 indicators	(TPACK-z.b.) The lesson serves to consolidate or deepen a subject content that has already been taught.
PCK	Content structuring	Transparency of learning objectives	4 indicators	(PCK-i1.c.) The teacher formulates clear content requirements for the video/animation.
		Focus on key aspects	5 indicators	(PCK-i2.a.) The teacher emphasizes the importance of key aspects of the content.
		Technical language	5 indicators	(PCK-i3.b.) The teacher defines technical terms to be used in the video/animation.
	Cognitive activation	Feedback	7 indicators	(PCK-i4.a.) The teacher defines the feedback criteria.
		Authenticity	2 indicators	(PCK-k1.a.) The video/animation is directed at a realistic audience.
		Relevance	3 indicators	(PCK-k2.b.) A relevant question or problem constitutes the basis of the video/animation creation.
		Transfer	6 indicators	(PCK-k3.d.) The teacher supports the students adequately with content-related issues.
TPK	Cognitive load	Use of DTM	5 indicators	(TPK-b1.c.) The teacher introduces the students to the mandatory functions of the software.
		Planning the SGEVA	6 indicators	(TPK-b2.e.) The teacher provides the students with a template (e.g., a storyboard) for planning their video.
		Group organization	4 indicators	(TPK-b3.b.) The teacher reduces the group size to small groups (3–4 students).
	Interactive learning		5 indicators	(TPK-u.a.) The teacher encourages students to engage in content discussions.
	Time allocation		3 indicators	(TPK-z.b.) Students are allocated sufficient time to plan and create their videos or animations (at least 40 min).

TABLE 3 Inter-rater reliability of the EnTPACK rubric at indicator and category level across the different data sets.

Data set	Cohen's Kappa	ICC
Planning	0.721	0.846
Implementation	0.723	0.848
Reflection	0.669	0.933
EnTPACK rubric	0.704	0.873

framework (Mishra and Koehler, 2006) as well as cognitively activating students and structuring the content (Steffensky and Neuhaus, 2018). The observable indicators were therefore firmly grounded in the related theory.

4.3. The EnTPACK rubric

Taking into account models of teacher professionalization (Baumert and Kunter, 2013; Stender et al., 2017) and recommendations from enacted TPACK research (Yeh et al., 2015; Pareto and Willermark, 2019), we designed the rubric for the combined measurement of lesson planning, implementation, and reflection using one instrument. This allows for a comparison of TPACK at the different parts of teacher professional development, providing comprehensive insights into PSSTs ability to effectively integrate technology in their teaching practice and helping to identify fractures in TPACK enactment. The identification of fractures in TPACK enactment also paves the way for the development of adapted content for teacher education that is capable of bridging the theory-practice gap. The rubric thus represents the first enacted TPACK instrument explicitly created for that purpose. Furthermore, it is situated on a specific use of DTM and science teaching approach under consideration of central contextual factors, enabling the framework's fuzzy facets to be operationalized more precisely. In contrast to previous approaches, we comprehensively and systematically reviewed the literature for this purpose, obtaining a sample of 59 articles to ensure the rubric was sufficiently grounded in the literature. In addition, we tested and inductively extended the rubric using a comprehensive case study, aligning the literature-based instrument with real-world practice.

To improve measurement objectivity, we provide clear descriptions of specific indicators for measuring the enacted TPACK of PSSTs. Utilizing 68 observable indicators for the grading of criteria renders the scoring process more objective than in existing instruments.

In further steps, the EnTPACK rubric will be used to examine how PSSTs apply their TPACK acquired (in a workshop) at the university in real classroom practice at the levels of lesson planning, implementation, and reflection. Several researchers highlight the scarcity of existing knowledge on the enactment of TPACK (Mouza, 2016; Ning et al., 2022), as well as a lack of studies examining TPACK enactment (Willermark, 2018). Accordingly, the developed rubric shall contribute to generate insights into this transfer process and to address the research gap.

The instrument's high degree of differentiation, its systematic grounding in theory and literature, the objective grading it

affords, and its comprehensive applicability to all three phases of teacher professionalization (lesson planning, implementation, and reflection) resulted in a highly complex instrument with many advantages—although some limitations must also be mentioned.

4.4. Limitations

Researchers face the challenge of designing TPACK instruments that are both valid and applicable to various settings, subjects, etc., (Young et al., 2012). The EnTPACK rubric presented in this paper represents a highly contextualized instrument based on an elaborate development process. The final instrument is optimized for its specific purpose, using SGEVA for teaching science content. However, this limits its applicability to the use of other DTMs. Transferable, as far as possible, are those indicators that do not explicitly refer to SGEVA. Accordingly, the validity of the instrument is restricted to the predefined conditions of use.

An upcoming task is to examine whether the rubric developed for the target group of PSSTs is also suitable for in-service science teachers. Also, developing and implementing the instrument requires a substantial investment of time and resources. Due to the complexity of the rubric a detailed coding manual as well as an elaborated rater training become indispensable.

Furthermore, there is a certain risk of bias in the conducted systematic reviews due to the restriction to the search engine google scholar and due to the limitation to 1,000 search results within the software "Publish or Perish." Despite the comprehensive size of google scholar, the inclusion of additional search engines and databases without a limitation of search results could have led to further results.

Finally, the EnTPACK rubric aims to address the use of SGEVA to learn and reinforce science content. The DPACK framework extends the professional knowledge of teachers by the aspect of critical reflection on DTM as part of a digital culture. This is not included in the our approach and should be considered more strongly in a future revision.

4.5. Implications for researchers and practitioners

The development of rubrics regarding TPACK is an elaborate theory-based process. The development process outlined in this paper can provide guidance to other researchers in developing instruments for other areas of pre- and in-service teachers' enacted TPACK. Given its applicability across lesson planning, implementation, and reflection, the instrument includes all relevant aspects of teachers' professional development.

It hence allows to survey the individual level of knowledge in planning, implementing, and reflecting on the use of DTM. In addition, however, the instrument also facilitates the identification of fractures between the levels. The identification of these fractures is crucial for the development of teaching materials that allow for a better fit between planning, implementation, and reflection at the university.

For practitioners, the instrument provides a basis, consistent with the current state of research, for how explainer videos can

be incorporated into science lessons and what criteria to consider when planning lessons and teaching with the SGEVA-approach. In addition, pre-service and in-service teachers can use the instrument to guide their self-reflection or peer-reflection with colleagues, thus supporting their own professionalization. To make this possible, the EnTPACK rubric will therefore be published in its entirety, as well as the corresponding coding manual. Currently, both the rubric and the coding manual are available on request from the corresponding author. In the future, these will be made publicly available via ResearchGate.

Author contributions

AA performed the systematic reviews and the development as well as validation of the instrument with supervision, support, and advice from HW. AA prepared the first draft of the manuscript. HW and SS revised the manuscript. All authors contributed equally to the conception and design of the study, final manuscript revision, read and approved the submitted version, and agree to be accountable for the content of the work.

Funding

The present work was funded by the German Federal Ministry of Education and Research (BMBF) (ref. 01JA2036). Responsibility

References

- Akaygun, S. (2016). Is the oxygen atom static or dynamic? The effect of generating animations on students' mental models of atomic structure. *Chem. Educ. Res. Pract.* 17, 788–807. doi: 10.1039/C6RP00067C
- Aktaş, I., and Özmen, H. (2020). Investigating the impact of TPACK development course on pre-service science teachers' performances. *Asia Pac. Educ. Rev.* 21, 667–682. doi: 10.1007/s12564-020-09653-x
- Aktaş, I., and Özmen, H. (2022). Assessing the performance of Turkish science pre-service teachers in a TPACK-practical course. *Educ. Inf. Technol.* 27, 3495–3528. doi: 10.1007/s10639-021-10757-z
- Al-Harhi, A. S. A., Campbell, C., and Karimi, A. (2018). Teachers' cloud-based learning designs: The development of a guiding rubric using the TPACK framework. *Comput. Sch.* 35, 134–151. doi: 10.1080/07380569.2018.1463033
- Andersen, M. F., and Munksby, N. (2018). Didactical design principles to apply when introducing student-generated digital multimodal representations in the science classroom. *Des. Learn.* 10, 112–122. doi: 10.16993/dfl.100
- Antony, M. K., and Paidi (2019). TPACK observation instrument: Development, validation, and reliability. *J. Phys. Conf. Ser.* 1241:012029. doi: 10.1088/1742-6596/1241/1/012029
- Ay, Y., Karadağ, E., and Acat, M. B. (2015). The technological pedagogical content knowledge-practical (TPACK-practical) model: Examination of its validity in the Turkish culture via structural equation modeling. *Comput. Educ.* 88, 97–108. doi: 10.1016/j.compedu.2015.04.017
- Baclay, L. (2020). Student-created videos as a performance task in science. *Int. J. Res. Stud. Educ.* 10, 1–11. doi: 10.5861/ijrse.2020.5925
- Baumert, J., and Kunter, M. (2013). "The COACTIV model of teachers' professional competence," in *Cognitive activation in the mathematics classroom and professional competence of teachers*, eds M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, and M. Neubrand (Boston, MA: Springer), 25–48. doi: 10.1007/978-1-4614-5149-5_2
- Becker, S., Meßinger-Kopelt, J., and Thyssen, C. (2020). *Digitale basiskompetenzen. Orientierungshilfe und praxisbeispiele für die universitäre lehramtsausbildung in den naturwissenschaften*. Hamburg: Joachim Herz Stiftung.
- Belski, L., and Belski, R. (2014). "Student-created dynamic (video) worked examples as a path to active learning," in *Proceedings of the 25th annual conference of the Australasian association for engineering education*, eds A. Brainbridge-Smith, Z. T. Qi,

and G. S. Gupta (Wellington, NZ: School of Engineering & Advanced Technology, Massey University), 1–9.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/feduc.2023.1190152/full#supplementary-material>

- and G. S. Gupta (Wellington, NZ: School of Engineering & Advanced Technology, Massey University), 1–9.
- Blacer-Bacolod, D. (2022). Student-generated videos using green screen technology in a biology class. *Int. J. Inf. Educ. Technol.* 12, 339–345. doi: 10.18178/ijiet.2022.12.4.1624
- Brantley-Dias, L., and Ertmer, P. (2014). Goldilocks and TPACK. *J. Res. Technol. Educ.* 46, 103–128. doi: 10.1080/15391523.2013.10782615
- Brianza, E., Schmid, M., Tondeur, J., and Petko, D. (2022). Situating TPACK: A systematic literature review of context as a domain of knowledge. *Contemp. Issues Technol. Teach. Educ.* 22, 707–753.
- Canbazoglu Bilici, S., Guzey, S. S., and Yamak, H. (2016). Assessing pre-service science teachers' technological pedagogical content knowledge (TPACK) through observations and lesson plans. *Res. Sci. Technol. Educ.* 34, 237–251. doi: 10.1080/02635143.2016.1144050
- Chai, C. S., Koh, J. H. L., and Tsai, C.-C. (2016). "A review of the quantitative measures of technological, pedagogical content knowledge (TPACK)," in *Handbook of technological pedagogical content knowledge (TPACK) for educators*, eds M. C. Herring, M. J. Koehler, and P. Mishra (New York, NY: Routledge), 87–106.
- Chi, M. T. H., and Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educ. Psychol.* 49, 219–243. doi: 10.1080/00461520.2014.965823
- Dewi, N. R., Rusilowati, A., Saptono, S., and Haryani, S. (2022). Project-based scaffolding TPACK model to improve learning design ability and TPACK of pre-service science teacher. *J. Pendidik. IPA Indones.* 11, 420–432. doi: 10.15294/jpii.v11i3.38566
- Döring, N. (2023). *Forschungsmethoden und evaluation in den sozial- und humanwissenschaften*, 6th Edn. Berlin: Springer.
- Drummond, A., and Sweeney, T. (2017). Can an objective measure of technological pedagogical content knowledge (TPACK) supplement existing TPACK measures? TPACK-deep and objective TPACK. *Br. J. Educ. Technol.* 48, 928–939. doi: 10.1111/bjet.12473
- Epps, B. S., Luo, T., and Salim Muljana, P. (2021). Lights, camera, activity! A systematic review of research on learner-generated videos. *J. Inf. Technol. Educ. Res.* 20, 405–427. doi: 10.28945/4874

- Farrokhnia, M., Meulenbroeks, R. F. G., and van Joolingen, W. R. (2020). Student-generated stop-motion animation in science classes: A systematic literature review. *J. Sci. Educ. Technol.* 29, 797–812. doi: 10.1007/s10956-020-09857-1
- Gallardo-Williams, M., Morsch, L. A., Paye, C., and Seery, M. K. (2020). Student-generated video in chemistry education. *Chem. Educ. Res. Pract.* 21, 488–495. doi: 10.1039/C9RP00182D
- Gusenbauer, M. (2019). Google Scholar to overshadow them all? Comparing the sizes of 12 academic search engines and bibliographic databases. *Scientometrics* 118, 177–214. doi: 10.1007/s11192-018-2958-5
- Harris, J. B., Grandgenett, N., and Hofer, M. (2010). “Testing a TPACK-based technology integration assessment rubric,” in *Proceedings of SITE 2010–Society for Information Technology and Teacher Education International Conference*, eds. D. Gibson, and B. Dodge (Chesapeake, VA: AACE), 3833–3840.
- Harris, J. B., Grandgenett, N., and Hofer, M. (2012). “Testing an instrument using structured interviews to assess experienced teachers’ TPACK,” in *Proceedings of SITE 2012–Society for Information Technology and Teacher Education International Conference*, ed. P. Resta (Austin, TX: AACE), 4696–4703.
- Hayati, N., Kadarohman, A., Sopandi, W., Martoprawiro, M. A., and Rochintaniawati, D. (2022). Chemistry teachers’ TPACK competence: Teacher perception and lesson plan analysis. *Tech. Soc. Sci. J.* 34, 237–247. doi: 10.47577/tssj.v34i1.7079
- He, J., and Huang, X. (2020). Using student-created videos as an assessment strategy in online team environments: A case study. *J. Educ. Multimedia Hypermedia* 29, 35–53.
- Hoban, G. (2020). “Slowmotion and blended media: Engaging students in a learning system when creating student-generated animations,” in *Learning from animations in science education. Innovating in semiotic and educational research*, ed. L. Unsworth (Cham: Springer), 193–208. doi: 10.1007/978-3-030-56047-8_8
- Hoban, G., and Nielsen, W. (2010). The 5 Rs: A new teaching approach to encourage slowmotions (student-generated animations) of science concepts. *Teach. Sci.* 56, 33–38.
- Hoban, G., and Nielsen, W. (2014). Creating a narrated stop-motion animation to explain science: The affordances of “slowmotion” for generating discussion. *Teach. Technol. Educ.* 42, 68–78. doi: 10.1016/j.tate.2014.04.007
- Hofer, M., Grandgenett, N., Harris, J., and Swan, K. (2011). “Testing a TPACK-based technology integration observation instrument,” in *Proceedings of the society for information technology and teacher education international conference 2011*, eds M. Koehler and P. Mishra (Nashville, TN: AACE), 4352–4359.
- Hoppe, H. U., Müller, M., Alissandrakis, A., Milrad, M., Schneegg, C., and Malzahn, N. (2016). “VC/DC”-video versus domain concepts in comments to learner-generated science videos,” in *Proceedings of the 24th international conference on computers in education (ICCE 2016)*, eds W. Chen, J.-C. Yang, S. Murthy, S. L. Wong, and S. Iyer (Mumbai: Asia Pacific Society for Computers in Education), 172–181.
- Hsu, Y.-S., Yeh, Y.-F., and Wu, H.-K. (2015). “The TPACK-P framework for science teachers in a practical teaching context,” in *Development of science teachers’ TPACK*, ed. Y.-S. Hsu (Singapore: Springer), 17–32. doi: 10.1007/978-981-287-441-2_2
- Huwer, J., Irion, T., Kuntze, S., Schaal, S., and Thyssen, C. (2019). Von TPACK zu DPACK—digitalisierung im unterricht erfordert mehr als technisches wissen. *MNU J.* 5, 358–364.
- Jacobs, B., and Cripps Clark, J. (2018). Create to critique: Animation creation as conceptual consolidation. *Teach. Sci.* 64, 26–36.
- Janssen, N., and Lazonder, A. W. (2016). Supporting pre-service teachers in designing technology-infused lesson plans: Support for technology-infused lessons. *J. Comput. Assist. Learn.* 32, 456–467. doi: 10.1111/jcal.12146
- Jen, T.-H., Yeh, Y.-F., Hsu, Y.-S., Wu, H.-K., and Chen, K.-M. (2016). Science teachers’ TPACK-practical: Standard-setting using an evidence-based approach. *Comput. Educ.* 95, 45–62. doi: 10.1016/j.compedu.2015.12.009
- Kafyulilo, A., Fisser, P., and Voogt, J. (2013). “TPACK development in teacher design teams: Assessing the teachers’ perceived and observed knowledge,” in *Proceedings of society for information technology and teacher education international conference 2013*, eds R. McBride and M. Searson (New Orleans, LA: AACE), 4698–4703.
- Kamp, B. L., and Deaton, C. C. M. (2013). Move, stop, learn: Illustrating mitosis through stop-motion animation. *Sci. Act. Classr. Proj. Curric. Ideas* 50, 146–153. doi: 10.1080/00368121.2013.851641
- Karakoyun, F., and Yapıcı, I. Ü (2018). Use of slowmotion in biology teaching. *Int. Educ. Stud.* 11:16. doi: 10.5539/ies.v11n10p16
- Kates, A. W., Wu, H., and Coryn, C. L. S. (2018). The effects of mobile phone use on academic performance: A meta-analysis. *Comput. Educ.* 127, 107–112. doi: 10.1016/j.compedu.2018.08.012
- Kavai, P., and Nyamupangedengu, E. (2021). “Exploring science teachers’ technological, pedagogical content knowledge at two secondary schools in Gauteng province of South Africa,” in *Proceedings of South Africa international conference on education*, eds M. Chitiyo, B. Seo, and U. Ogbonnaya (Winston-Salem, NC: AARF), 122–132.
- Keast, S., and Cooper, R. (2011). “Developing the knowledge base of preservice science teachers: Starting the path towards expertise using slowmotion,” in *The Professional knowledge base of science teaching*, eds D. Corrigan, J. Dillon, and R. Gunstone (Dordrecht: Springer Netherlands), 259–277. doi: 10.1007/978-90-481-3927-9_15
- Kidman, G., Keast, S., and Cooper, R. (2012). “Understanding pre-service teacher conceptual change through slowmotion animation,” in *Proceedings of the 2nd international STEM in education conference*, ed. S. Yu (Beijing: Beijing Normal University), 194–203.
- KMK (2017). *Bildung in der digitalen Welt. Strategie der Kultusministerkonferenz*. Available online at: https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2018/Strategie_Bildung_in_der_digitalen_Welt_idf_vom_07.12.2017.pdf (accessed January 20, 2023).
- Koo, T. K., and Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* 15, 155–163. doi: 10.1016/j.jcm.2016.02.012
- Kuckartz, U., and Rädiker, S. (2022). *Qualitative inhaltsanalyse: Methoden, praxis, computerunterstützung: Grundlagentexte methoden*, 5th Edn. Basel: Beltz Juventa.
- Labudde, P., and Möller, K. (2012). Stichwort: Naturwissenschaftlicher unterricht. *Z. Für Erzieh.* 15, 11–36. doi: 10.1007/s11618-012-0257-0
- Landis, J. R., and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics* 33:159. doi: 10.2307/2529310
- Lawrie, G., and Bartle, E. (2013). Chemistry Vlogs: A vehicle for student-generated representations and explanations to scaffold their understanding of structure - property relationships. *Int. J. Innov. Sci. Math. Educ.* 21, 27–45.
- Ling Koh, J. H., Chai, C. S., and Tay, L. Y. (2014). TPACK-in-Action: Unpacking the contextual influences of teachers’ construction of technological pedagogical content knowledge (TPACK). *Comput. Educ.* 78, 20–29. doi: 10.1016/j.compedu.2014.04.022
- Max, A., Lukas, S., and Weitzel, H. (2022). The relationship between self-assessment and performance in learning TPACK: Are self-assessments a good way to support preservice teachers’ learning? *J. Comput. Assist. Learn.* 38, 1160–1172. doi: 10.1111/jcal.12674
- Mills, R., Tomas, L., and Lewthwaite, B. (2019). The impact of student-constructed animation on middle school students’ learning about plate tectonics. *J. Sci. Educ. Technol.* 28, 165–177. doi: 10.1007/s10956-018-9755-z
- Mishra, P., and Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teach. Coll. Rec.* 108, 1017–1054. doi: 10.1111/j.1467-9620.2006.00684.x
- Molenda, M. (2015). In search of the elusive ADDIE model. *Perform. Improv.* 54, 40–42. doi: 10.1002/pfi.21461
- Mourlam, D., and Bleecker, H. (2017). “Early career teacher candidate TPACK development: Implementation of a learning activity types short course,” in *Proceedings of Society for Information Technology & Teacher Education International Conference*, eds P. Resta, and S. Smith (Austin, TX: AACE), 2404–2409.
- Mourlam, D., Chesnut, S., and Bleecker, H. (2021). Exploring preservice teacher self-reported and enacted TPACK after participating in a learning activity types short course. *Australas. J. Educ. Technol.* 37, 152–169. doi: 10.14742/ajet.6310
- Mouza, C. (2016). “Developing and assessing TPACK among pre-service teachers. A synthesis of research,” in *Handbook of technological pedagogical content knowledge (TPACK) for educators*, eds M. C. Herring, M. J. Koehler, and P. Mishra (New York, NY: Routledge), 169–190.
- Nielsen, W., Turney, A., Georgiou, H., and Jones, P. (2020). “Creating a digital explanation in preservice teacher education: Scientific knowledge represented in a digital artefact,” in *Learning from animations in science education. Innovating in semiotic and educational research*, ed. L. Unsworth (Cham: Springer International Publishing), 229–248. doi: 10.1007/978-3-030-56047-8_10
- Ning, Y., Zhou, Y., Wijaya, T. T., and Chen, J. (2022). Teacher education interventions on teacher TPACK: A meta-analysis study. *Sustainability* 14:11791. doi: 10.3390/su141811791
- Njiku, J., Mutarutinya, V., and Maniraho, J. F. (2020). Developing technological pedagogical content knowledge survey items: A review of literature. *J. Digit. Learn. Teach. Educ.* 36, 150–165. doi: 10.1080/21532974.2020.1724840
- Orraryd, D. (2021). *Making science come alive: Student-generated stop-motion animations in science education*. Linköping: Linköping University Electronic Press, doi: 10.3384/diss.diva-173038
- Özdilek, Z., and Robeck, E. C. (2019). Enabling pre-service chemistry teachers’ development of technological, pedagogical, and content knowledge (TPACK) through case-based lesson planning. *J. Uludag Univ. Fac. Educ.* 32, 207–225. doi: 10.19171/uefad.533233
- Özdilek, Z., and Ugur, B. (2019). Experiences of pre-school teacher candidates on the development and application of slowmotion in science education. *J. Educ. Learn.* 8, 152–167. doi: 10.5539/jel.v8n5p152
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Syst. Rev.* 10:89. doi: 10.1186/s13643-021-01626-4
- Paidi, Hapsari, N., Subali, B., and Astuti, F. E. C. (2020). “Teaching performance of high school biology teachers in applying TPACK: A descriptive study,” in *Proceedings*

- of the international conference on educational research and innovation (ICERI 2019), (Yogyakarta, IN), doi: 10.2991/assehr.k.200204.018
- Palmgren-Neuvonen, L., and Korkeamäki, R.-L. (2014). Group interaction of primary-aged students in the context of a learner-generated digital video production. *Learn. Cult. Soc. Interact.* 3, 1–14. doi: 10.1016/j.lcsi.2013.11.001
- Palmgren-Neuvonen, L., and Korkeamäki, R.-L. (2015). Teacher as an orchestrator of collaborative planning in learner-generated video production. *Learn. Cult. Soc. Interact.* 7, 1–11. doi: 10.1016/j.lcsi.2015.09.002
- Palmgren-Neuvonen, L., Korkeamäki, R.-L., and Littleton, K. (2017). Intercreating in the context of learner-generated DV production. *Think. Ski. Creat.* 26, 13–23. doi: 10.1016/j.tsc.2017.04.005
- Pareto, L., and Willermark, S. (2019). TPACK in situ: A design-based approach supporting professional development in practice. *J. Educ. Comput. Res.* 57, 1186–1226. doi: 10.1177/0735633118783180
- Petko, D. (2020). *Einführung in die mediendidaktik: Lehren und Lernen mit digitalen medien*, 2nd Edn. Basel: Beltz.
- Pirhonen, J., and Rasi, P. (2016). Student-generated instructional videos facilitate learning through positive emotions. *J. Biol. Educ.* 51, 215–227. doi: 10.1080/00219266.2016.1200647
- Popham, W. J. (1997). What's wrong-and what's right-with rubrics. *Educ. Leadersh.* 55, 72–75.
- Potter, M., Tuck, K. L., Robinson, A. J., Richardson, P. W., and Grieve, A. (2021). MediaLab: Video as a multivalent tool for science teaching and learning. *Media Pract. Educ.* 22, 23–37. doi: 10.1080/25741136.2021.1832766
- Redecker, C. (2017). *European framework for the digital competence of educators: DigCompEdu*. Luxembourg: Publications Office of the European Union, doi: 10.2760/178382
- Reyna, J., and Meier, P. (2018a). Learner-generated digital media (LGDM) as an assessment tool in tertiary science education: A review of literature. *IAFOR J. Educ.* 6, 93–109. doi: 10.22492/ije.6.3.06
- Reyna, J., and Meier, P. (2018b). Using the learner-generated digital media (LGDM) framework in tertiary science education: A pilot study. *Educ. Sci.* 8:106. doi: 10.3390/educsci8030106
- Ribosa, J., and Duran, D. (2022). Do students learn what they teach when generating teaching materials for others? A meta-analysis through the lens of learning by teaching. *Educ. Res. Rev.* 37:100475. doi: 10.1016/j.edurev.2022.100475
- Robeck, E., and Sharma, A. (2011). Student-generated multimedia as a means to meaningful science education. *J. Int. Soc. Teach. Educ.* 15, 50–59.
- Rosenberg, J. M., and Koehler, M. J. (2015). Context and technological pedagogical content knowledge (TPACK): A systematic review. *J. Res. Technol. Educ.* 47, 186–210. doi: 10.1080/15391523.2015.1052663
- Seibert, J., Kay, C. W. M., and Huwer, J. (2019). EXPLAINistry: Creating documentation, explanations, and animated visualizations of chemistry experiments supported by information and communication technology to help school students understand molecular-level interactions. *J. Chem. Educ.* 96, 2503–2509. doi: 10.1021/acs.jchemed.8b00819
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educ. Res.* 15, 4–14. doi: 10.3102/0013189X015002004
- So, H.-J., and Kim, B. (2009). Learning about problem based learning: Student teachers integrating technology, pedagogy and content knowledge. *Australas. J. Educ. Technol.* 25, 101–116. doi: 10.14742/ajet.1183
- Steffensky, M., and Neuhaus, B. J. (2018). “Unterrichtsqualität im naturwissenschaftlichen unterricht,” in *Theorien in der naturwissenschaftsdidaktischen forschung*, eds D. Krüger, I. Parchmann, and H. Schecker (Berlin: Springer), 299–313. doi: 10.1007/978-3-662-56320-5_18
- Stender, A., Brückmann, M., and Neumann, K. (2017). Transformation of topic-specific professional knowledge into personal pedagogical content knowledge through lesson planning. *Int. J. Sci. Educ.* 39, 1690–1714. doi: 10.1080/09500693.2017.1351645
- Sweller, J., Van Merriënboer, J. J. G., and Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educ. Psychol. Rev.* 10, 251–296. doi: 10.1023/A:1022193728205
- Thyssen, C., Huwer, J., Irion, T., and Schaal, S. (2023). From TPACK to DPACK: The “digitality-related pedagogical and content knowledge”-model in STEM-education. *Educ. Sci.* 13:769. doi: 10.3390/educsci13080769
- Tondeur, J., Aesaert, K., Prestridge, S., and Consuegra, E. (2018). A multilevel analysis of what matters in the training of pre-service teacher's ICT competencies. *Comput. Educ.* 122, 32–42. doi: 10.1016/j.compedu.2018.03.002
- Tondeur, J., van Braak, J., Sang, G., Voogt, J., Fisser, P., and Ottenbreit-Leftwich, A. (2012). Preparing pre-service teachers to integrate technology in education: A synthesis of qualitative evidence. *Comput. Educ.* 59, 134–144. doi: 10.1016/j.compedu.2011.10.009
- Unal Çoban, G., Kocagül, M., Akpınar, E., and Baran, B. (2022). Examining science teachers' performances at planning geology lesson through TPACK-based argumentation practices. *Kuramsal Eğitim Bilim Derg.* 15, 907–937. doi: 10.30831/akukey.1110596
- von Kotzebue, L. (2022). Beliefs, self-reported or performance-assessed TPACK: What can predict the quality of technology-enhanced biology lesson plans? *J. Sci. Educ. Technol.* 31, 570–582. doi: 10.1007/s10956-022-09974-z
- Voogt, J., Fisser, P., Pareja Roblin, N., Tondeur, J., and van Braak, J. (2012). Technological pedagogical content knowledge - a review of the literature. *J. Comput. Assist. Learn.* 29, 109–121. doi: 10.1111/j.1365-2729.2012.00487.x
- Wang, W., Schmidt-Crawford, D., and Jin, Y. (2018). Preservice teachers' TPACK development: A review of literature. *J. Digit. Learn. Teach. Educ.* 34, 234–258. doi: 10.1080/21532974.2018.1498039
- White, P. J., Tytler, R., and Nielsen, W. (2020). “Animation construction as cross-modal translation in senior biology,” in *Learning from animations in science education. Innovating in semiotic and educational research*, ed. L. Unsworth (Cham: Springer), 209–228. doi: 10.1007/978-3-030-56047-8_9
- Willermark, S. (2018). Technological pedagogical and content knowledge: A review of empirical studies published from 2011 to 2016. *J. Educ. Comput. Res.* 56, 315–343. doi: 10.1177/0735633117713114
- Willmott, C. (2014). Teaching bioethics via the production of student-generated videos. *J. Biol. Educ.* 49, 1–12. doi: 10.1080/00219266.2014.897640
- Wu, H.-K., and Puntambekar, S. (2012). Pedagogical affordances of multiple external representations in scientific processes. *J. Sci. Educ. Technol.* 21, 754–767. doi: 10.1007/s10956-011-9363-7
- Wu, Y.-T. (2013). Research trends in technological pedagogical content knowledge (TPACK) research: A review of empirical studies published in selected journals from 2002 to 2011. *Br. J. Educ. Technol.* 44, E73–E76. doi: 10.1111/j.1467-8535.2012.01349.x
- Yaseen, Z. (2016). *Student-generated animations and the teaching and learning of chemistry*. Dissertation. Sydney, NSW: University of Technology.
- Yaseen, Z. (2018). Using student-generated animations: The challenge of dynamic chemical models in states of matter and the invisibility of the particles. *Chem. Educ. Res. Pract.* 19, 1166–1185. doi: 10.1039/C8RP00136G
- Yaseen, Z., and Aubusson, P. (2018). Exploring student-generated animations, combined with a representational pedagogy, as a tool for learning in chemistry. *Res. Sci. Educ.* 50, 529–548. doi: 10.1007/s11165-018-9700-4
- Yeh, Y.-F., Chien, S.-P., Wu, H.-K., and Hsu, Y.-S. (2015). “Rubrics of TPACK-P for teaching science with ICTs,” in *Development of science teachers' TPACK: East Asian practices*, ed. Y.-S. Hsu (Singapore: Springer), 54–70.
- Young, J. R., Young, J. L., and Shaker, Z. (2012). Technological pedagogical content knowledge (TPACK) literature using confidence intervals. *TechTrends* 56, 25–33. doi: 10.1007/s11528-012-0600-6