OPEN ACCESS
Jurnal Ilmiah Ilmu Terapan Universitas Jambi
Vol. 8 No. 2. December 2024, pp. 577-592 DOI. 10.22437/jiituj.v8i2.33958
Research Article

SCALE FOR PRE-SERVICE SCIENCE TEACHERS' KNOWLEDGE IN DESIGNING INCLUSIVE LEARNING: CONSTRUCTION, VALIDATION, AND RELIABILITY

Jamil Suprihatiningrum[1,*], Astuti Wijayanti[2], Firman Walidi Salam[1], Muhammad Aditya Hidayah[3]

[1] Faculty of Education and Teacher Training, Universitas Islam Negeri Sunan Kalijaga Yogyakarta, Yogyakarta, Indonesia
[2] Faculty of Education and Teacher Training, Universitas Sarjanawiyata Tamansiswa Yogyakarta, Yogyakarta, Indonesia
[3] Faculty of Science, Prince of Songkla University, Songkhla, Thailand
Corresponding author email: jamil.suprihatiningrum@uin-suka.ac.id

Article Info
Recieved: Jun 02, 2024
Revised: Sep 09, 2024
Accepted: Oct 20, 2024
OnlineVersion: Oct 28, 2024

Abstract
Scale development is essential in testing constructs and variables in social research, including education. Scales that meet the rules of validity and reliability are developed using standardized procedures. This research focuses on scale production to create reliable and valid construct measures to assess teacher knowledge in designing inclusive learning, using three frameworks, namely Technological Pedagogical and Content Knowledge, Backward Design, and Universal Design for Learning. Research on this issue involves measuring organizational and psychological constructs, which present unique challenges because they are generally unobservable, very abstract, often complex, and may consist of several different components. As a result of this complexity, developing this scale was challenging. The production of this scale followed Boateng et al., and the field test involved 259 pre-service teachers. The Rasch Rating Scale Model analysis was used to test the scale and produced 18 items that have high reliability (0.96). This scale can be used nationally to assess teachers' ability to implement inclusive education. The scale also applies as a teacher self-assessment tool for developing their competence.

Keywords: Backward Design, Inclusive Pre-Service Teachers, Scale Validation, TPACK, UDL

INTRODUCTION

Pre-Service Teacher Training (PTT) is the first form of professional training for future teachers to enter the teaching profession at a specified level of education (UIS-UNESCO, n.d.), which usually consists of theoretical knowledge of teaching and practicum. In Indonesia, the PTT can only be organized by a university with a faculty of education and teacher training (LPTK) selected by the Ministry of Education and Culture. In 2018, 43 LPTKs (out of 400) met the requirements to become PTT organizers, and from this list, only 29 LPTKs were permitted to offer PTT (Yusrina et al., 2022).

The quality of training provided by the PTT influences the practice, effectiveness, and quality of the graduate students, ultimately impacting the quality of learning and education (Tasdemir et al., 2020). Therefore, LPTK must offer qualified PTT to prepare competent future teachers (Padagas, 2019) and build future teachers' identities (Torres-Cladera et al., 2021). To generate potential and skills in students, they must be able to express their thinking as creatively as possible (Iqbal et al., 2023). In addition, new graduates experience a major transition from the academic environment to the world of work, demanding adaptation from the role of student to worker, which requires a wide array of skills that may have been less honed during college (Kusuma, 2020; Sari, Omeiza, & Mwakifuna, 2023; Nugroho et al., 2024).

Providing a quality PTT program can help pre-service teachers succeed in preparing learning that is adaptive to changing times while improving the quality of learning and education for all (EFA). Introducing the concept of inclusive education in a PTT program is a way to realize the EFA. Indonesia has demonstrated its commitment to inclusive education by issuing various regulations, such as the latest Minister of National Education Regulation No. 48 of 2023, which regulates reasonable accommodation for students with disabilities and the establishment of disability service units in each educational unit. However, inclusive education implementation in schools is not evenly distributed, and no national teacher training program focuses on inclusive education.

A preliminary study of the PTT curriculum shows that only one course leads to inclusive learning and is included in elective courses, namely Introduction to Education for Children with Special Needs, and a load of two credits (Ministry of Education and Culture, 2022). This indicates that inclusive education is not mandatory in the PPT Program, while the future challenges are that teachers will face students with more diversity. Massouti (2021) mentioned that the PTT Program must also be criticized regarding curriculum structure, primarily how the program design and requirements will better support future teachers' knowledge and practice around inclusive learning. Future teachers need to be prepared for the dynamics of student diversity, as teachers play a crucial role in protecting students' rights to learn (Wang & Shih, 2022). Lack of training in inclusive learning (Forlin et al., 2009; Florian et al., 2010; Binti M & Adeshina, 2024; Syahputra & Edwards, 2024) causes newly graduated teachers to face instructional challenges in supporting diverse learners and contributes to the reproduction of unfair practices in schools (López-Torrijo & Mengual-Andrés, 2015; Subban & Mahlo, 2016). These challenges may ultimately prevent the admission of students who have been historically marginalized based on ability, language, faith, culture, ethnicity, or socioeconomic background (Mitchell, 2017).

On the other hand, several researchers (Goldem & Dore, 2001; Wulff et al., 2004; Reynolds & Kearns, 2017) reported that primary and secondary education teachers received little or no study about learning design from their lecturers while at university. Lecturers need to assist teachers in developing their lesson design practices (Koh et al., 2015). There are various frameworks for designing learning, including Backward Design (BD). Studies by Daugherty (2006); Wood (2009); Trigwell (2010); Ebert-May et al. (2011); Baker (2014); and Dolan and Collins (2015) conclude that there have been efforts to equip lecturers about the use of BD theory and processes and its impact on student-centered learning practices. BD has become an effective planning strategy for achieving student-centered, outcome-based learning. Wiggins and McTighe (2005) proposed using a BD process through three stages that focus on 1) identifying the desired results, 2) analyzing various data sources, and 3) determining the appropriate action plan to achieve the desired results that have been determined previously. Learning with BD requires educators to go out of the box (out of the "standards") that are usually used so that educators can have more freedom in determining what students need and want to achieve through authentic mastery of the material (McTighe & Thomas, 2003). Such an approach allows educators to view concepts through various lenses that reveal how students understand essential and non-essential lesson content, leading to learning for all children.

Besides BD, inclusive learning can be designed using the Universal Design for Learning (UDL) Framework. By clarifying willingness to learn and providing a flexible learning environment, the UDL framework can identify and overcome student learning difficulties (King-Sears, 2009; Burton et al., 2010; Spencer, 2011; King-Sears et al., 2015; Suwarni, 2021). UDL can help school personnel design learning because it directs the way to align learning objectives and practices to create a positive relationship between student interest and involvement in learning (Smith, 2012). UDL allows teachers to approach each student individually (whether with or without disabilities), offering pathways to access the curriculum, implementing alternative instruction, adjusting the learning pace, and providing various ways to demonstrate student understanding (Mason & Orkwis, 2005). UDL can also be used to design inclusive learning that includes new technologies (McMahon & Walker, 2019). Despite the potential of the UDL framework to support pre-service teachers in inclusive classrooms, there is a limited amount of research examining its implementation in real-world teaching contexts during teacher preparation programs (Mackey et al., 2023).

Besides, teacher knowledge regarding pedagogy, content, and technology, known as the Technological Pedagogical and Content Knowledge (TPACK) framework, has been discussed recently in Indonesia. This framework is widely applied in Indonesia's PTT Programs. Despite the many approaches to TPACK development, empirical studies of Information and Communication Technology (ICT) practices in schools over the last decade show teachers using technology primarily for content presentation (Ruthven et al., 2004; Smeets, 2005; Hayes, 2007; Lim & Chai, 2008; Gao et al., 2009; Ward & Parr, 2010; Player-Koro, 2013; Uluyol and Sahin, 2016). ICT professional development programs focusing on TPACK also report inconclusive results about changes in teacher ICT practice. Teachers who underwent a TPACK development program that focused on the integration of specific technological devices demonstrated improvements in technological skills after attending the program, but not all expressed confidence in integrating devices in a way that supported student-centered learning (Niess 2007; Blau et al., 2014). In content-focused TPACK development programs where student-centered use of ICT is explicitly modeled for specific curriculum topics, reported adoption of this strategy is also low, with teachers citing contextual barriers in schools (for example Jimoyiannis (2010)). TPACK development programs focusing on specific pedagogical approaches appear to be more successful in helping teachers articulate student-centered ICT lesson design (Doering et al., 2014; Jaipal-Jamani & Figg, 2015; Koh et al., 2017; Ozden et al., 2024).

Literature studies show that no research applies BD, UDL, and TPACK in one framework for designing learning, especially those aimed at inclusive settings. Therefore, creating a scale to measure the pre-service teachers' knowledge in designing inclusive learning is essential. This research will have positive implications for the PTT Program and recommendations for LPTK, teachers, and related policymakers regarding how inclusive learning is designed and implemented in classes with diverse students.

RESEARCH METHOD

The development of this scale followed the steps suggested by Boateng et al. (2018), which consists of three main phases, namely "item development, scale development, and scale evaluation" (p. 2), and is described into nine steps as seen in Figure 1. Item development was executed by identifying the domains and the existing literature on BD, UDL, and TPACK. The initial items were then validated by experts using a review form. Scale development was initiated by pre-testing the items to the target population and gathering their feedback, and was continued with the survey administration to the pre-service teachers. The survey's data was then analyzed in two steps: item reduction and factor analysis. The data were then evaluated using three tests: dimensionality, item function, and assessment category function. Before the study was conducted, the Institute for Research and Community Service, UIN Sunan Kalijaga Yogyakarta, Number B-293/Un.02/L3/TL/12/2023, granted the ethical clearance.

RESULTS AND DISCUSSION

Phase 1 – Item Development

Step 1 – Identification of the Domain(s) and Item Generation

This stage began with identifying the domains that will be measured, namely pre-service teachers' knowledge in designing inclusive learning, which is defined as the information possessed by pre-service teachers in designing learning, including creating lesson plans, teaching materials, media, and assessments referring to three frameworks, namely BD, UDL, and TPACK. After conducting an online search through repositories and sources (e.g., PubMed, ERIC, Google Scholar, Publish or Perish), no scale was developed for measuring pre-service teachers' knowledge using those three frameworks. The preliminary conceptual definition was then determined as follows.

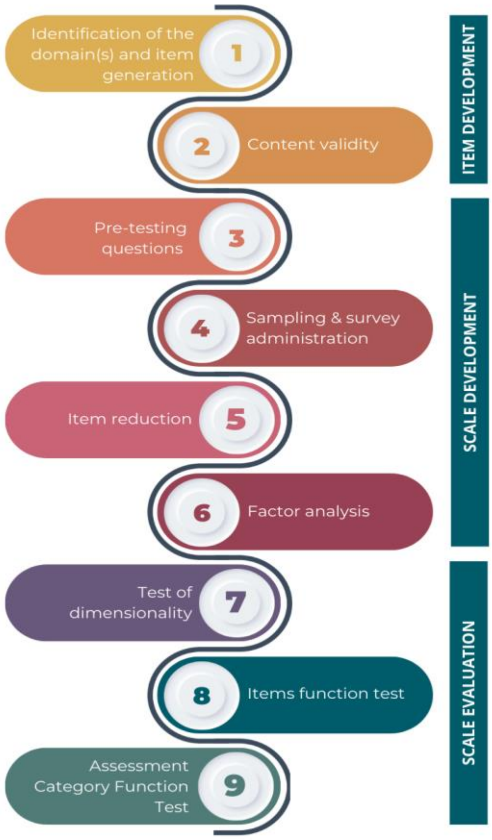

Figure 1. Scale development procedure

1. BD requires educators to start with a list of essential questions gathered from the previously mentioned out-of-the-box steps that students must answer at the end of the learning topic (Jones et al., 2009). With this aim, the teacher designs assessments aligned with the concept/lesson material, providing a framework for delivering lessons (learning activities). The desired outcome is an engaging lesson that aligns with the material and assessment. Ideally, learning activities reflect the skills and goals the learning topic assesses. Both novice and senior teachers often forget what it is like to be a novice learner, so this BD process allows for the type of planning that ensures learning is aligned with the curriculum and focuses on students in depth (covering the three domains of affective, cognitive and psychomotor) and does not expand during learning activities ongoing (Au, 2007). Reynolds and Kearns (2017) suggested that other benefits of BD include priority delivery of material to students, effective time management, reduced anxiety, increased creativity in idea solutions, and engagement of students by offering more excellent feedback about their understanding.
2. Research on PTT Programs that use the UDL framework in design learning is minimal. Some studies appear to include UDL as a component (Allday et al., 2013; Marino et al., 2009; Van Laarhoven et al., 2012) but do not include the implemented or focused intervention. One study that used UDL as an intervention is Lowrey et al. (2019). They involved eight pre-service teachers and intervened with two phases of student teaching experiences: one traditional and one after receiving UDL professional development. The research shows that minimal UDL teaching results in less inclusive learning planning. Israel et al. (2014) offered recommendations for pre-service teacher preparation content courses, in which pre-service teachers should have a basic understanding of the UDL framework as a prerequisite for UDL implementation.

   Since all effective teaching aims to promote authentic learning, understanding UDL's unique contribution to effective teaching is critical. Mason and Orkwis (2005) stated UDL: assumes diverse learners, where students will learn at different levels and each have strengths and weaknesses in specific areas; depends on a standard curriculum, presented flexibly to engage all learners; proposes high expectations for all (including for SWD) so that they can be assessed under the same standards; and "Inclusive UDL by design" means all instructional components are prepared and available before the lesson begins.

3. The discourse regarding TPACK is still happening, especially in the Indonesian context, giving rise to many teachers' professional development programs focusing on TPACK, including developing teacher TPACK for student-centered use of ICT. For example, Walker et al. (2012) and Koh et al. (2017) studied the development of teachers' TPACK by prioritizing student-centered pedagogical dimensions such as problem-based learning and meaningful learning with ICT; Tai (2015) and Jimoyiannis (2010) modeled TPACK-specific curricula for English and Mathematics, respectively. In addition to focusing on specific technological tools, curriculum, or pedagogy, some programs seek to engender teacher TPACK through collaborative design, problem-solving, and reflection (e.g., Boschman et al. (2015)) and programs focused on developing teacher TPACK to support initiatives one-to-one laptops in their schools (Blau et al., 2014).

The final conceptual definition regarding the pre-service teachers' knowledge in designing inclusive learning in this research is the integrated pedagogical, content, and technological knowledge possessed by pre-service teachers to design curriculum and learning experiences to meet the needs of diverse children, starting with determining the end goals and then work backward to develop teaching materials and learning activities that meet those goals.

After that, we carried out item generation using two methods, namely deductive and inductive methods. Deductive methods are carried out in two ways: searching the literature regarding BD, UDL, and TPACK and assessing similar scales developed. The literature search results regarding BD, UDL, and TPACK are presented in Table 1, and the existing scale is in Table 2. The literature review provided a theoretical basis for defining the domain, operationalizing the conceptual definition from theory into a measurable and observable scale. Scales defined by theoretical underpinnings are better placed to make specific, pragmatic decisions about domains because the construct will be based on accumulated knowledge about existing items. Items generation also involved inductive methods through discussions by the developer and interviews with three potential scale users. Combining both methods was considered the best practice for defining domains, identifying questions to assess them (Boateng et al., 2018), as well as producing more comprehensive scales (Clark & Watson, 1995; Loevinger, 1957).

Table 1. Literature search on BD, UDL, and TPACK

Author(s)	Year	Title
Lumberas, Jr. R., & Rupley, W. H.	2020	Pre-service teachers' application of Understanding by Design in lesson planning
Joubert et. al.	2020	Lesson study in a blended approach to support isolated teachers in teaching with technology
Hills et. al.	2020	Using anticipated learning outcomes for Backward Design of a molecular cell biology course-based undergraduate research experience
Schmidt et. al.	2009	Technological Pedagogical Content Knowledge (TPACK)
Chai et. al.	2011	Modelling primary school pre-service teachers' Technological Pedagogical Content Knowledge (TPACK) for meaningful learning with information and communication technology (ICT)
Valtonen et. al.	2017	TPACK updated to measure pre-service teachers' twenty-first century skills
Scott et. al.	2015	UDL in Online College Coursework: Insights of Infusion and Educator Preparedness
Benton-Borghi	2016	Universal design for learning (UDL) infused technological pedagogical content knowledge (TPACK) model prepares efficacious 21st-century teachers
Kennete, L. N., & Wilson, N. A.	2019	Universal Design for Learning (UDL): student and faculty perceptions
Cash et. al.	2021	Distance Educators Attitudes and Actions towards Inclusive Teaching Practices

Table 2. Existing Research on the Scale of BD, UDL, and TPACK

Author(s)	Year	Title
Kiray.	2016	Development of a TPACK self-efficacy scale for preservice science teachers
Sang et al.	2016	Validation and profile of Chinese pre-service teachers' technological pedagogical content knowledge scale
Al-Harthi et al.	2018	Teachers' cloud-based learning designs: The development of a guiding rubric using the TPACK framework
Schmid et al.	2020	Developing a short assessment instrument for Technological Pedagogical Content Knowledge (TPACK.xs) and comparing the factor structure of an integrative and a transformative model
Adipat	2021	Developing Technological Pedagogical Content Knowledge (TPACK) through Technology-Enhanced Content and Language-Integrated Learning (T-CLIL) Instruction

Several deliberations were made in item generation, including the form of the item used, which was "statement," not "question; the item wording, which was simple, unambiguous, contained one criterion, and was not biased by social identity, gender, race, religion, or economic status. At this stage, it produced 35 items.

This scale used a five-point response because, according to Boateng et al. (2018), unipolar items would reflect the relative degree of single-item response quality, for example, "can't do it at all" to "very good at it." A three-point response was also not chosen because, according to Krosnick and Presser (2009), a response with only two to three points has lower reliability than a scale with five to seven points.

Step 2: Content Validity

The 35 items were evaluated by experts (to ensure each item has content relevance, representativeness, and technical quality) and the target population (to represent the target population's experience) (Boateng et al., 2018). Seven experts reviewed the items and dropped five points from 35 to 30. Their comments are shown in Table 3.

Table 3. Experts' reviews

Reviewer	Comments
Expert 1	Aspects of PK can be added with points regarding teacher skills in uncovering student characteristics and put this in item number one. If this scale combines three frameworks, then it is better if each item contains these three aspects. UDL aspects need to be included in all statement items. Item No. 19, "My ability to design and implement online learning (both synchronous/asynchronous) that allows students to build new knowledge and skills", can be connected to the UDL concept related to how teachers provide various ways so that children can understand the material provided through the technology. Item No. 20, "My ability to provide additional time to students who need it, for submitting assignments and exams in various formats (paper and digital)", can be related to how teachers provide various ways so that children are able to demonstrate their learning results.
Expert 2	Questionnaires can be made better by using Google Forms so that they can be filled in directly and can be analysed according to the researcher's needs.
Expert 3	In general, all questions are good and easy for respondents to understand.
Expert 4	The items in each aspect include explanations, and the narrative in each statement is good and complete so that respondents can understand the context of the statement clearly.
Expert 5	The questionnaire is good, there are explanations of real examples such as TikTok, Facebook, etc. So, the reader understands what the researcher means.
Expert 6	Overall, all questions were good and easy for respondents to understand.
Expert 7	Item No. 1, pedagogical competency, might include 1 or 2 learning theories because sometimes we know the theory but forget the name or term.

Phase 2: Scale Development

Step 3: Pre-testing Questions

The pre-testing questions step was used to check the extent to which the questions reflect the domain being studied and how the answers to the questions were asked to produce valid measurements. Pre-testing questions were carried out on ten target populations to answer 30 items; then, they clarified their experience in fulfilling the scale through unstructured interviews. The reviewed target population is in Table 4. Only seven out of ten respondents provided their reviews, and three others stated that the scale was feasible.

Step 4: Survey Administration

The revisions were made based on the experts and the target population, producing 30 items. The next step was administering the survey to pre-service teachers by distributing the Google Forms link via WhatsApp to the heads of study programs organizing PTT in the regions of DI Yogyakarta, Central Java, South Sulawesi, DI Aceh, and South Kalimantan to be forwarded to students. They were selected conveniently and randomly, and as many as 259 students were willing and agreed to fill out the scale. The profile of respondents is known based on the distribution of respondents involved in this study (Pasulu et al., 2023).

Step 5: Item Reduction Analysis

Step 5 was used to identify items that are not or at least related to the domain under study to be deleted or modified so that only functional and internally consistent items were included using Item Response Theory (IRT), as suggested by Boateng et al. (2018).

Table 4. Target Population Reviews

Target Population	Comments
1	No. 29 Point 2 can clarify what to make time for.
2	In general, the questions are easy to understand, but there are some that could be clarified further.
3	The points on the designed scale include things that need to be measured.
4	The items can be understood and contain clear meaning
5	No. 30 can be added to learning experiences in everyday life.
6	Overall, it is good but it would be nice to improve the words that are more relevant.
7	No. 1 can be explained about what learning theory is being studied No. 2 can be shortened No. 3 can be separated per indicator No. 4 scoring and rubrics are replaced with assessment guidelines No. 6 and 7 do not need the word “for example” No. 11, the sequence of concepts is replaced with a learning flow. No. 29 can be replaced to create a conducive and meaningful learning environment.

Step 6: Factor Analysis

This step tested the scale’s number of factors (i.e., variables). Removing unnecessary factors will reduce the number of dimensions in the data, which can reduce the amount of residual data that IRT cannot explain. This analysis can also be used to identify cases to be analysed and to limit the factors to be analysed in IRT.

Phase 3: Scale Evaluation

The data collected was then tested on the models (i.e., Rasch Rating Scale Model (RSM) and Partial Credit Model (PCM)) to find which one suits the data obtained (Bond et al., 2020; Boone & Staver, 2020). The criteria used to determine the most appropriate model for data analysis are the lowest values of Akaike’s Information Criterion (AIC), Bayesian’s Information Criterion (BIC), and Corrected Akaike’s Information Criterion (CAIC). Table 5 shows that the RSM was the most suitable compared to the PCM.

Table 5. Model Comparison

Model	Deviance	AIC	BIC	CAIC	Parameters	N
PCM	7697	7879	8203	8294	91	259
RSM	7769	7815	7897	7920	23	259

RSM analysis was carried out using Winsteps (version 3.73) and Jamovi (version 2.4) software with the SnowRMM module (Linacre, 2012; Seol, 2023; The jamovi project, 2023). In RSM, three stages were carried out: testing the scale dimensionality, testing the item's function, and testing the assessment scale (Bond et al., 2020; Boone & Staver, 2020).

Step 7: Dimensionality Test

The dimensionality test was carried out in two stages: the Exploratory Factor Analysis (EFA) and the misfit order test. The EFA test was carried out using Jamovi software (Versi 2.4; The jamovi project, 2023) by considering the root mean square error of approximation (RMSEA) value below 0.80; the Tucker-Lewis Index (TLI) is close to 0.90 or above, and the p-value model exceeds 0.050 (Hu & Bentler, 1998; Whittaker & Schumacker, 2022). In addition, the data used in the EFA must comply with the Barlett and Kaiser-Meyer-Olkin (KMO) assumption test, with the Barlett test criteria having a p-value below 0.05 and KMO having an adequacy value above 0.60 (Pett et al., 2003; Rasli, 2006). The misfit order test criteria in this study were carried out using Winstep software by considering the value of Outfit MNSQ (0.6-1.4) for the rating scale model (Bond et al., 2020; Boone & Staver, 2020).

Table 6. Summary of Exploratory Factor Analysis

Factors	Variables	Loadings	Variance (%)
1	PCK	2.129	11.83
2	IEK	1.873	10.40
3	PK	1.872	10.40
4	CK	1.703	9.46
5	TK	1.671	9.28
6	TPACK	1.356	7.53
7	PTK	1.339	7.44
8	TCK	0.346	1.92

The results of the EFA analysis show that of the 30 proposed items, only 18 items were accepted, namely PK2, PK3, PK5, TK1, TK2, CK1, CK2, TCK3, PCK1, PCK2, PCK4, PTK3, TPACK1, TPACK3, TPACK4, IEK2, IEK3, IEK4. Table 6 shows that the approximate factor contained in the scale is eight factors that correspond with the designed factors. The EFA shows that the PCK has the highest variance value, and the TCK has the lowest variance value. Table 7 shows that the designed model's value is sufficient for the required criteria with an RMSEA value of 0.00, TLI of 1.02, and p-value of 0.927. In the Barlett test, the p-value was found to be below 0.001, which shows that the sample used was normally distributed, whereas, in the KMO test, it was found that the average value for each item was 0.931, which shows that the number of samples used was sufficient (Pett et al., 2003; Rasli, 2006).

Table 7. Model Fit

	RMSEA 90% CI				Model Test		
RMSEA	Lower	Upper	TLI	BIC	χ^2	df	p
0.00	0.00	0.0141	1.02	-180	25.3	37	0.927

The misfit order test was carried out with Winsteps software (version 3.73) to determine the difficulty level of the proposed items (Linacre, 2012). All items fit with the predetermined criteria. Once the model is fit, a reliability test will be carried out based on the valid items. The reliability test result shows that Cronbach's α value is at 0.929, which indicates that the scale used in this research is very reliable. Table 8 shows that the MNSQ outfit values obtained range from 0.74 to 1.35, sorted from the most straightforward items to the most difficult. The EFA and misfit order tests concluded that the items used are unidimensional.

Table 8. Summary of Misfit Order and Item Difficulty

Item	Measure	PT. Measure Corr.	Outfit MNSQ	Perceived Difficulties
PK2	-2.45	0,66	0,95	Most Difficult item
TCK3	-2.71	0,72	1,08	↑
TPACK4	-2.71	0,71	0,86	
PK3	-2.84	0,63	1,02	
TPACK1	-3.27	0,62	1,20	
TPACK3	-3.30	0,67	1,19	
PCK2	-3.37	0,73	0,79	
PCK1	-3.41	0,73	0,76	
CK1	-3.46	0,74	0,74	
TK2	-3.47	0,58	1,35	
PK5	-3.50	0,71	0,91	
IEK3	-3.53	0,68	1,03	
CK2	-3.63	0,72	0,78	
IEK4	-3.89	0,68	0,99	
PCK4	-4.03	0,68	0,98	
IEK2	-4.08	0,66	1,00	
PTK3	-4.56	0,63	1,14	↓
TK1	-4.59	0,65	1,00	Easiest item

Step 8: Items Function Test

Testing the items function was carried out using a Wright map using Jamovi software (Version 2.4; Seol, 2023: The Jamovi project, 2023) to explain a direct comparison of the distribution of respondents' abilities with the distribution of item difficulty levels on the same logit scale (Bond et al., 2020; Boone & Staver, 2020). Figure 2 shows that the distribution of respondents' abilities was on the left side of the map, and the distribution of item difficulty levels was on the right side. Respondents at the top of the map quickly agreed with items on the scale. In contrast, respondents at the bottom of the map found it challenging to agree with items on the scale. Items at the top of the map were challenging to approve, while items at the bottom were easy to approve.

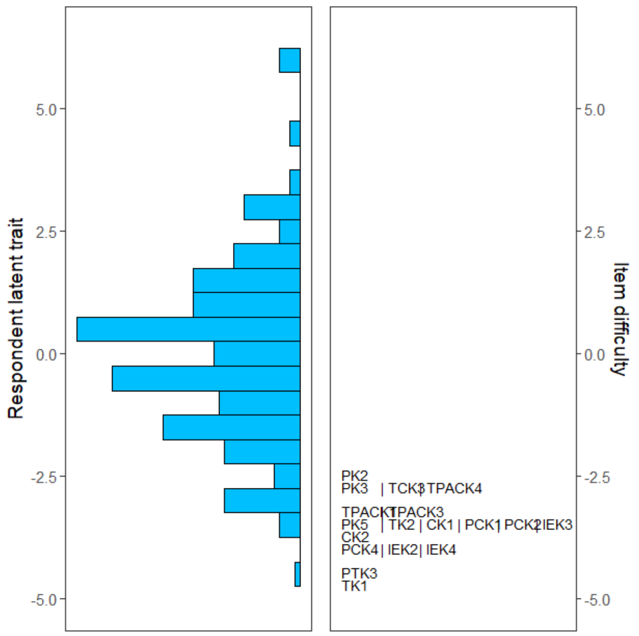

Figure 2. Wright map BD-TPACK-UDL

Figure 2 demonstrates that the distribution of respondents is at logit -5 to 6, but the distribution was concentrated at logit -2.5 to 2.5, with the mean being at logit 0.0. This map also demonstrates that the difficulty levels of items have a narrower distribution than the distribution of respondents' abilities, namely logit -5 to logit -2.5, with the mean logit -3. Thus, it can be said that the items on the scale are too easy for the average respondent to agree with, judging from the distance between the mean respondent's ability and the level of item difficulty (3 logits). Figure 2 also shows a long lag between respondents who were at logit 6 at the top of the map and logit 5, but at the bottom, the lag for respondents was not too long. On the right side, there was no gap between items, which indicates that the difficulty level of the items was well distributed.

The next item function test used reliability and separation of respondents and items using Winsteps software (version 3.73; Linacre, 2012). Fisher (2007) stated that the reliability of items and respondents is above 0.67, while separating items and respondents is above 2.0. The results of data analysis show that the reliability and item separation are 0.96 (excellent) and 4.71 (very good). Furthermore, the reliability and separation of respondents was 0.91 (very good) and 3.23 (good). Thus, the designed scale can define the difficulty level of items.

Step 9: Assessment Category Function Test

The assessment category function test aims to determine how well the answer categories are used in the scale. Linacre (1999) created procedures for testing the function of assessment categories, including the minimum number of observations for each observation is 10. Table 9 shows that each category has met the minimum number of observations except for category 1 (bad), which has three observations. This table shows that the highest number of observations is category 4 (very good), with a total of 2470 observations, and the lowest acceptable number of observations is category 2 (low), with a total of 83 observations. In addition, the scale needs to be tested with the criteria that the MNSQ outfit value must be below 2.00. Table 9 shows that each category has an MNSQ outfit value below 2.00, demonstrating good function.

Table 9. Calibration Scaling Analysis

Category Label	Observed Count (%)	Observed Average	Infit MNSQ	Outfit MNSQ	Andrich Threshold	Category Measure
1 (very poor)	3 (0%)	0,62	1,68	1,78	None	(-5,43)
2 (poor)	83 (2%)	0,38	1,46	1,43	-4,22	-3,32
3 (good)	1318(28%)	1,07	0,96	0,95	-2,41	-0,45
4 (very good)	2470 (53%)	3,04	0,87	0,87	1,47	3,32
5 (excellent)	788 (17%)	4,83	1,02	1,00	5,15	(6,27)

The next step was to test the Andrich Threshold in each category with the criterion that the logit difference between one category and another is more than 1.40 logits (Linacre, 1999, 2012). Table 9 shows the difference between categories 1 and 2 is 4.22 logits, between categories 2 and 3 is 1.81 logits, between categories 3 and 4 is 3.88 logits, and between categories 4 and 5 is 3.68 logits. Thus, it can be seen that each category clearly distinguishes respondents' abilities.

The next test was to use the category probability curve to determine the distribution of categories on the scale. In RSM, each item was modelled to have the same response structure, but this does not mean that respondents judge that each item is the same (Boone & Staver, 2020; Linacre, 2012). A good category is a category that has a peak with a probability above 50% (0.50) and forms a clear hill (Bond et al., 2020; Boone & Staver, 2020). Figure 3 illustrates an x-axis showing the respondent's ability, while the y-axis shows the probability of selecting a category. Thus, as the respondent's logit increases, the probability of being in a better category will increase. Figure 3 shows each category has a peak above 0.50, and each category forms a clear hill.

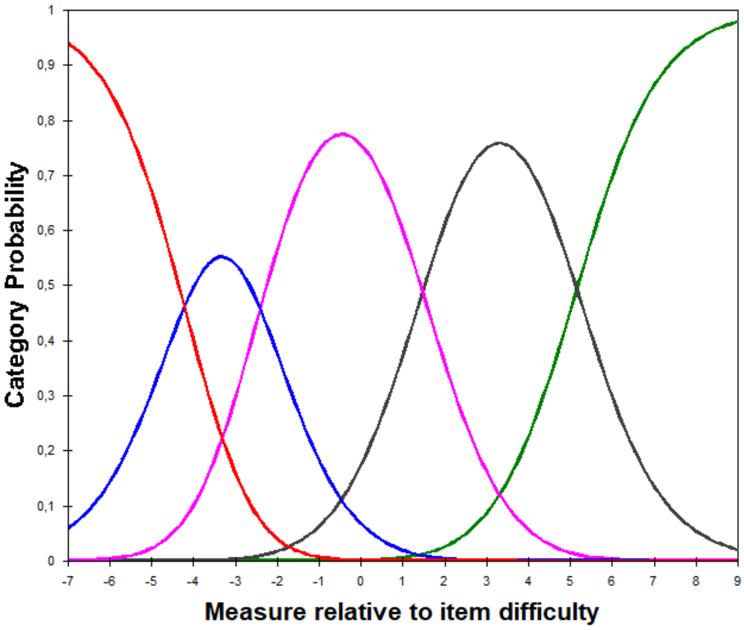

Figure 3. Category Probability Curve

Step 10: Test Item Bias

The final step in testing items was to test for differences in answer characteristics in two or more different groups of respondents. This was done so that the designed scale does not provide unilateral benefits to specific groups of respondents (Bond et al., 2020; Boone & Staver, 2020). Testing item bias on the RSM can be done using differential item functioning (DIF) by looking at the chi-square's probability value. The item bias test in this study was carried out with Winsteps software (version 3.73). The probability value criterion is considered significant if the value is below 0.05 (Linacre, 2012). In this study, the respondent groups used were gender (DIF 1), field of study (DIF 2), experience teaching students with disabilities (DIF 3), university origin (DIF 4), and university pursuing professional teacher education (DIF 5).

Table 10. Differential Item Functioning (DIF)

Item	Prob. DIF 1	Prob. DIF 2	Prob. DIF 3	Prob. DIF 4	Prob DIF 5
PK 2	0,391	0,754	1,000	1,000	0,527
PK 3	1,000	0,522	0,670	0,090	0,138
PK 5	0,675	0,627	0,667	0,812	0,452
TK 1	0,760	0,250	0,212	0,120	1,000
TK 2	0,054	0,382	0,168	0,086	0,001*
CK 1	0,893	0,330	1,000	0,027*	0,070
CK 2	0,590	0,420	0,237	0,175	1,000
TCK 3	0,210	0,002*	0,343	1,000	0,016*
PCK 1	0,424	0,018*	0,556	0,675	0,747
PCK 2	1,000	0,446	0,272	0,700	1,000
PCK 4	0,004*	0,114	0,292	0,774	0,814
PTK 3	0,508	1,000	0,087	1,000	0,684
TPACK 1	0,544	0,570	0,583	0,074	0,340
TPACK 3	0,065	0,025*	0,689	0,360	0,663
TPACK 4	0,210	1,000	0,558	0,479	0,149
IEK 2	0,888	0,239	0,868	0,560	0,246
IEK 3	0,569	0,404	0,602	0,743	0,651
IEK 4	0,182	0,214	0,677	1,000	0,665

Table 10 shows that overall, the respondent group has different answer characteristics when viewed from the probability value below 1.000. Still, this difference can be said to be not significant. The DIF 1 group has a probability value ranging from 0.004 to 1.000 with question items with significant differences in answers to PCK 4 questions. The DIF 2 group has a probability value ranging from 0.002 to 1.000 with question items with significant answer differences in TCK 3, PCK 1, and TPACK 3. The DIF 3 group has a probability value range from 0.168 to 1.000, which does not show a significant difference in answers. The DIF 4 group has a probability value ranging from 0.027 to 1.000, significantly differing in answers to CK 1. The DIF 5 group has a probability value range from 0.001 to 1.000 with questions that significantly differ in answers to the items TK 2 and TCK 3. Thus, the designed scale needs to be re-evaluated by removing items indicated as biased by changing the items or re-writing them.

Literature studies have found that there is no scale to measure PTT students' knowledge in designing inclusive learning. This study confirms that teachers can use the BD-UDL-TPACK framework to help create inclusive learning. It also assists them in evaluating their knowledge and abilities in working with students with disabilities, which integrates with technology. Therefore, LPTK and teacher training universities can widely use this scale. The results of measuring pre-service teacher knowledge in designing inclusive learning through this scale can be used to evaluate the PTT curriculum as well as to review policies for ongoing teacher career development programs in Indonesia. In addition, when using this scale, users should consider potential confounding variables that may influence the results, such as differences in teaching experience, training, or exposure to inclusive education practices, so that the results of the scale can be explored more broadly to identify areas where pre-service teachers need additional support in designing inclusive learning environments.

CONCLUSION

This research succeeded in testing a scale to measure the knowledge of PTT students to design inclusive learning based on the BD-UDL-TPACK framework using Rasch rating scale model analysis. In the unidimensionality test using EFA, eight factors were extracted that corresponded to the variables designed in the scale, with PCK having the highest variance (11,83%), and TCK having the lowest variance (1,92%). The instrument consists of 18 items that met the expected model fit criteria, with the RMSEA value being 0.00, the TLI value being 1.02, and the p-value being 0.927. The Barlett and KMO tests show that the data obtained through the scale is normally distributed with a sufficient sample size for analysis using EFA. The MNSQ outfit score meets the recommended criteria for testing respondents' abilities. The Cronbach's α value is at 0.929, showing that the scale is reliable.

This study has several limitations. *First*, the dimensionality test only used EFA. For future research, the test dimensionality can be done using CFA. *Second*, this study applied a Rasch rating scale model, which assumes the same response structure for each item in the category function test. Future research is expected to test the scale using a graded response model or partial credit model. *Third*, the bias test in this study used DIF from the standard Rasch rating scale model; future research is expected to test DIF using the Rasch mixture model. *Fourth*, methodologically, this study did not consider contextual factors, such as the specific characteristics of the PTT, the cultural and socioeconomic background of the participants, or the broader educational context. Therefore, further research is recommended to take this into account. *Fifth*, a larger sample size will increase the generalisability of the findings, so it is recommended that future researchers use larger samples and longitudinal studies so they can evaluate changes in pre-service teacher knowledge over time.

ACKNOWLEDGMENTS

This research project was supported by a grant from the Research Center of UIN Sunan Kalijaga Yogyakarta, Number 75.8/2023.

AUTHOR CONTRIBUTIONS

Conceptualization, J.S.; Methodology, J.S., A.W.; Software, F.W.; Validation, J.S., A.W., F.W., M.A.H.; Formal Analysis, J.S.; Investigation, J.S. A.W.; Resources, J.S., F.W., M.A.H.; Data Curation, F.W.; Writing – Original Draft Preparation, J.S.; Writing – Review & Editing, J.S., A.W., M.A.H.; Visualization, F.W.; Supervision, J.S.; Project Administration, J.S.; Funding Acquisition, J.S.

CONFLICTS OF INTEREST

The author(s) declare no conflict of interest.

REFERENCES

Allday, R. A., Neilsen-Gatti, S., & Hudson, T. M. (2013). Preparation for inclusion in teacher education pre-service curricula. *Teacher Education and Special Education*, *36*, 298–311. https://doi.org/10.1177/0888406413497485

Au, W. (2007). High stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher*, *36*(5), 258-267. https://doi.org/10.3102/0013189X07306523

Baker, L. A. e. a. (2014). Collateral scholars' collaborative new faculty workshop: Professional development for new chemistry faculty and initial assessment of its efficacy. *Journal of Chemistry Education*, *91*(11), 1874-1881. https://doi.org/10.1021/ed500547n

Binti M, M., & Adeshina, A. N. G. (2024). Exploring the effectiveness of the learning cycle model in enhancing mathematics learning for students. *Interval: Indonesian Journal of Mathematical Education*, *2*(2), 45-51. https://doi.org/10.37251/ijome.v2i2.1144.

Blau, I., Peled, Y., & Nusan, A. (2014). Technological, pedagogical and content knowledge in one-to-one classroom: Teachers developing ''digital wisdom''. *Interactive Learning Environments*, *13*, 1–16. https://doi.org/10.1080/10494820.2014.978792

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Front Public Health*, *6*, 149. https://doi.org/10.3389/fpubh.2018.00149

Bond, T. G., Yan, Z., & Heene, M. (2020). *Applying the rasch model: Fundamental measurement in the human sciences*. https://doi.org/10.4324/9780429030499

Boone, W. J., & Staver, J. R. (2020). *Advances in Rasch Analyses in the Human Sciences*. https://doi.org/10.1007/978-3-030-43420-5

Boschman, F., McKenney, S., & Voogt, J. (2015). Exploring teachers' use of TPACK in design talk: The collaborative design of technology-rich early literacy activities. *Computers & Education*, *82*, 250–262. https://doi.org/10.1016/j.compedu.2014.11.010

Burton, M., Krall, J., Devlin, C., Hannon, M., James, K., Rice, K., Blak, T., & Strang, A. (2010). *Universal Design for Learning in BC*. Special Education Technology - British Columbia.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7*(3), 309–319. https://doi.org/10.1037/1040-3590.7.3.309

Daugherty, K. K. (2006). Backward course design: Making the end the beginning. *American Journal of Pharmaceutical Education*, *70*(6), 1-5.

Doering, A., Koseoglu, S., Scharber, C., Henrickson, J., & Lanegran, D. (2014). Technology integration in K–12 geography education using TPACK as a conceptual model. *Journal of Geography*, *113*(6), 223–237. https://doi.org/10.1080/00221341.2014.896393

Dolan, E. L., & Collins, J. P. (2015). We must teach more effectively: Here are four ways to get started. *Molecular Biology of the Cell*, *25*(12), 2151-2155. https://doi.org/10.1091/mbc.E13-11-0675.

Ebert-May, D., Derting, T. L., Hodder, J., Momsen, J. L., Long, T. M., & Jardeleza, S. E. (2011). What we say is not what we do: Effective evaluation of faculty professional development programs. *BioScience*, *61*(7), 550-558. https://doi.org/10.1525/bio.2011.61.7.9

Florian, L., Young, K., & Rouse, M. (2010). Preparing teachers for inclusive and diverse educational environments: Studying curricular reform in an initial teacher education course. *International Journal of Inclusive Education*, *14*(7), 709-722. https://doi.org/10.1080/13603111003778536

Forlin, C., Loreman, T., Sharma, U., & Earle, C. (2009). Demographic differences in changing pre-service teachers' attitudes, sentiments and concerns about inclusive education. *International Journal of Inclusive Education*, *13*(2), 195-209. https://doi.org/10.1080/13603110701365356

Gao, P., Choy, D., Wong, A. F. L., & Wu, J. (2009). Developing a better understanding of technology based pedagogy. *Australasian Journal of Educational Technology*, *25*(5), 714–730. https://doi.org/10.14742/ajet.1117

Goldem, C. M., & Dore, T. M. (2001). *At Cross Purposes: What the Experiences of Today's Doctoral Students Reveal about Doctoral Education: A Survey Initiated by the Pew Charitable Trusts*.

Hayes, D. N. A. (2007). ICT and learning: Lessons from Australian classrooms. *Computers & Education*, *49*(2), 385–395. https://doi.org/10.1016/j.compedu.2005.09.003

Iqbal, M., Farida, L. Z. N., & Win, K. T. (2023). The influence of student attitudes on learning achievement. *Jurnal Ilmiah Ilmu Terapan Universitas Jambi*, *7*(2), 92-98. https://doi.org/10.22437/jiituj.v7i2.26697

Israel, M., Ribuffo, C., & Smith, S. (2014). *Universal design for learning innovation configuration: Recommendations for teacher preparation and professional development*. University of Florida, Collaboration for Effective Educator, Development, Accountability, and Reform Center.

Jaipal-Jamani, K., & Figg, C. (2015). A case study of a TPACK-based approach to teacher professional development: Teaching science with blogs. *Contemporary Issues in Technology and Teacher Education*, *15*(2), 161–200.

Jimoyiannis, A. (2010). Designing and implementing an integrated technological pedagogical science knowledge framework for science teachers professional development. *Computers & Education*, *55*(3), 1259–1269. https://doi.org/10.1016/j.compedu.2010.05.022

Jones, K. A., Vermette, P. J., & Jones, J. L. (2009). An integration of backwards planning unit design with the two-step lesson planning framework. *Education*, *130*(2), 357-360.

King-Sears, M. (2009). Universal design for learning: Technology and pedagogy. *Learning Disability Quarterly*, *32*(4), 199-201. https://doi.org/10.2307/27740372

King-Sears, M. E., Johnson, T. M., Berkeley, S., Weiss, M. P., Peters-Burton, E. E., Evmenova, A. S., Menditto, A., & Hursh, J. C. (2015). An exploratory study of universal design for teaching chemistry to students with and without disabilities. *Learning Disability Quarterly*, *38*(2), 84–96. https://doi.org/10.1177/0731948714564575

Koh, J. H. L., Chai, C. S., Hong, H.-Y., & Tsai, C.-C. (2015). A survey to examine teachers' perceptions of design dispositions, lesson design practices, and their relationships with technological pedagogical content knowledge (TPACK). *Asia-Pacific Journal of Teacher Education*, *43*(5), 378-391. https://doi.org/10.1080/1359866X.2014.941280

Koh, J. H. L., Chai, C. S., & Lim, W. Y. (2017). Teacher professional development for TPACK-21CL Effects on teacher ICT integration and student outcomes. *Journal of Educational Computing Research*, *55*(2), 172–196. https://doi.org/10.1177/0735633116656848

Krosnick, J. A., & Presser, S. (2009). Question and questionnaire design. In W. JD & M. PV (Eds.), *Handbook of Survey Research* (pp. 263–314). Elsevier.

Kusuma, R. S. (2020). Improving students' basic asking skills by using the discovery learning model. *Tekno - Pedagogi : Jurnal Teknologi Pendidikan*, *10*(2), 8-13. https://doi.org/10.22437/teknopedagogi.v10i2.32743.

Lim, C. P., & Chai, C. S. (2008). Teachers' pedagogical beliefs and their planning and conduct of computer-mediated classroom lessons. *British Journal of Educational Technology*, *39*(5), 807–828. https://doi.org/10.1111/j.1467-8535.2007.00774.x

Linacre, J. M. (1999). Investigating rating scale category utility. *Journal of outcome measurement*, *3*(2).

Loevinger, J. (1957). Objective Tests as Instruments of Psychological Theory. *Psychological Reports*, *3*(3), 635-694. https://doi.org/10.2466/pr0.1957.3.3.635

Lopez-Torrijo, M., & Mengual-Andres, S. (2015). An attack on inclusive education in secondary education. Limitations in initial teacher training in Spain. *Journal of New Approaches in Educational Research*, *4*(1), 9-17. https://doi.org/10.7821/naer.2015.1.100

Lowrey, K. A., Classen, A., & Sylvest, A. (2019). Exploring ways to support preservice teachers' use of UDL in planning and instruction. *Journal of Educational Research and Practice*, *9*(1), 261–281. https://doi.org/10.5590/JERAP.2019.09.1.19

Mackey, M., Drew, S. V., Nicoll-Senft, J., & Jacobson, L. (2023). Advancing a theory of change in a collaborative teacher education program innovation through universal design for learning. *Social Sciences & Humanities Open*, *7*(1), 100468. https://doi.org/https://doi.org/10.1016/j.ssaho.2023.100468

Marino, M., Sameshima, P., & Beecher, C. (2009). Enhancing TPACK with assistive technology: Promoting inclusive practices in pre-service teacher education. *Contemporary Issues in Technology and Teacher Education*, *9*, 186-207.

Mason, C., & Orkwis, R. (2005). Universal Design for Learning in Today's Diverse Schools. In C. f. E. Children (Ed.), *Universal design for learning : A guide for teachers and education professionals*. Council for Exceptional Children, Pearson Custom [and] Merrill Education.

Massouti, A. (2021). Pre-service teachers' perspectives on their preparation for inclusive teaching: Implications for organizational change in teacher education. *The Canadian Journal for the Scholarship of Teaching and Learning*, *12*(1). https://doi.org/10.5206/cjsotl-rcacea.2021.1.10611

McMahon, D. D., & Walker, Z. (2019). Leveraging emerging technology to design an inclusive future with Universal Design for Learning. *Center for Educational Policy Studies Journal*, *9*(3), 75-93. https://doi.org/10.26529/cepsj.63976

McTighe, J., & Thomas, R. S. (2003). Backward design for forward action. *Educational Leadership*, *60*(5), 52-55.

Ministry of Education and Culture. (2022). *Arah kebijakan PPG Prajabatan 2022 (translation: Pre-service teacher training education 2022 policy direction)*. https://ppg.kemdikbud.go.id/prajab/detail-program-kurikulum

Mitchell, D. R. (2017). *Diversities in education: Effective ways to reach all learners*. Routledge.

Niess, M. L. (2007). Developing teacher's TPCK for teaching mathematics with spreadsheets. *Technology and Teacher Education Annual*, *18*(4), 2238–2245.

Nugroho, N. E., Irianto, J., & Suryanto, S. (2024). A systematic review of indonesian higher education students' and graduates' work readiness. *Jurnal Ilmiah Ilmu Terapan Universitas Jambi*, *8*(1), 350-363. https://doi.org/10.22437/jiituj.v8i1.33073

Ozden, S. Y., Yang, H., Wen, H., & Shinas, V. H. (2024). Reflections from a teacher education course built on the TPACK framework: Examining the impact of the technology integration planning cycle on teacher candidates' TPACK development and practice. *Social Sciences & Humanities Open*, *9*, 100869. https://doi.org/https://doi.org/10.1016/j.ssaho.2024.100869

Padagas, R. C. (2019). Pre-service teachers' competencies in a work-based learning environment. *African Educational Research Journal*, *7*(3), 130-142. https://doi.org/10.30918/AERJ.73.18.076

Pasulu, Y. T., Budiyanto, A. B., & Astuti, R. A. (2023). Evaluation of the level of patient satisfaction with pharmacy services at dr. J.P. Wanane, Sorong district. *Jurnal Ilmiah Ilmu Terapan Universitas Jambi*, *7*(2), 175-183. https://doi.org/10.22437/jiituj.v7i2.30113

Player-Koro, C. (2013). Hype, hope and ICT in teacher education: A Bernsteinian perspective. *Learning, Media and Technology*, *38*(1), 26–40. https://doi.org/10.1080/13562517.2015.1069266

Reynolds, H. L., & Kearns, K. D. (2017). A planning tool for incorporating backward design, active learning, and authentic assessment in the college classroom. *College Teaching*, *65*(1), 17-27. https://doi.org/10.1080/87567555.2016.1222575

Ruthven, K., Hennessy, S., & Brindley, S. (2004). Teacher representations of the successful use of computer-based tools and resources in secondary-school english, mathematics and science. *Teaching and Teacher Education*, *20*(3), 259–275. https://doi.org/10.1016/j.tate.2004.02.002

Sari, R., Omeiza, I. I., & Mwakifuna, M. A. (2023). The influence of number dice games in improving early childhood mathematical logic in early childhood education. *Interval: Indonesian Journal of Mathematical Education*, *1*(2), 61-66. https://doi.org/10.37251/ijome.v1i2.776.

Smeets, E. (2005). Does ICT contribute to powerful learning environments in primary education? *Computers & Education*, *44*(3), 343–355. https://doi.org/10.1016/j.compedu.2004.04.003

Smith, F. G. (2012). Analyzing a college course that adheres to the Universal Design for Learning (UDL) Framework. *Journal of Scholarship of Teaching and Learning*, *12*(3), 31-61.

Spencer, S. A. (2011). Universal design for learning: Assistance for teachers in today's inclusive classrooms. *Interdisciplinary Journal of Teaching and Learning*, *1*(1), 10-22.

Subban, P., & Mahlo, D. (2016). 'My attitude, my responsibility' investigating the attitudes and intentions of pre-service teachers toward inclusive education between teacher preparation cohorts in Melbourne and Pretoria. *International Journal of Inclusive Education*, *21*(4), 441-461. https://doi.org/10.1080/13603116.2016.1197322

Suwarni, R. (2021). Analysis the process of observing class iv students in thematic learning in primary schools. *Tekno - Pedagogi : Jurnal Teknologi Pendidikan*, *11*(1), 26-32. https://doi.org/10.22437/teknopedagogi.v11i1.32717.

Syahputra, G., & Edwards, A. J. (2024). Transforming history teaching: Using adobe photoshop e-posters to teach the battle of november 10, 1945 . *Journal of Educational Technology and Learning Creativity*, *2*(1), 29-40. https://doi.org/10.37251/jetlc.v2i1.980.

Tai, S.-J. D. (2015). From TPACK-in-action workshops to classrooms: CALL competency developed and integrated. *Language Learning & Technology*, *19*(1), 139–164.

Tasdemir, M. Z., Iqbal, M. Z., & Asghar, M. Z. (2020). A study of the significant factors affecting pre-service teacher education in Turkey. *Bulletin of Education and Research*, *42*(1), 79-100.

https://doi.org/https://www.proquest.com/docview/2435558348/6393575D187E463APQ/6?accountid=11365

Torres-Cladera, G., SimÃ³-GilN., Domingo-PeÃ±afielL., & Amat-Castells, V. (2021). Building professional identity during pre-service teacher education. *Center for Educational Policy Studies Journal*, *11*(3), 35-54. https://doi.org/10.26529/cepsj.107036

Trigwell, K. (2010). Promoting Effective Student Learning in Higher Education. In P. Peterson, E. Baker, & B. McGaw (Eds.), *International Encyclopedia of Education* (pp. 461-466). Elsevier.

UIS-UNESCO. (n.d.). Pre-service teacher training. In *Glossary*. http://uis.unesco.org/en/glossary-term/pre-service-teacher-training

Uluyol, C., & Sahin, S. (2016). Elementary school teachers' ICT use in the classroom and their motivators for using ICT. *British Journal of Educational Technology*, *47*(1), 65–75. https://doi.org/10.1111/bjet.12220

Van Laarhoven, T., Munk, D. D., Chandler, L. K., Zurita, L., & Lynch, L. (2012). Integrating assistive technology into teacher education programs: Trials, tribulations, and lessons learned. *Assistive Technology Outcomes and Benefits*, *8*, 32-47.

Walker, A., Recker, M., Ye, L., Robertshaw, M. B., Sellers, L., & Leary, H. (2012). Comparing technology- related teacher professional development designs: A multilevel study of teacher and student impacts. *Educational Technology Research and Development*, *60*(3), 421–444.

Wang, R.-J., & Shih, Y.-H. (2022). Improving the quality of teacher education for sustainable development of Taiwan's education system: A systematic review on the research issues of teacher education after the implementation of 12-year national basic education. *Frontiers in Psychology*, *13*(921839), 1-12. https://doi.org/10.3389/fpsyg.2022.921839

Ward, L., & Parr, J. M. (2010). Revisiting and reframing use: Implications for the integration of ICT. *Computers & Education*, *54*(1), 113–122. https://doi.org/10.1016/j.compedu.2009.07.011

Wiggins, G., & McTighe, J. (2005). *Understanding by Design* (2nd ed.). ASCD.

Wood, W. B. (2009). Innovations in teaching undergraduate biology and why we need them. *Annual Review of Cell and Developmental Biology*, *25*(5), 93-112. https://doi.org/10.1146/annurev.cellbio.24.110707.175306

Wulff, D. H., Austin, A. E., Hyquist, J. D., & Sprague, J. (2004). The Development of Graduate Students as Teaching Scholars: A Four-Year Longitudinal Study. In D. H. Wulff & A. E. Austin (Eds.), *Paths to the Professoriate: Strategies for Enriching the Preparation of Future Faculty* (pp. 46-74).

Yusrina, A., Alifia, U., Revina, S., Putri Pramana, R., & Bima, L. (2022). *Is the game worth the candle? Examining the effectiveness of initial teacher education in Indonesia* RISE Working Paper Series.

Appendix

The scale of Pre-service Science Teachers' Knowledge of Inclusive Learning Design based on TPACK-UDL-BD

Instructions

Give a self-assessment of your own knowledge by placing a tick (√) in the selected column. Your knowledge is assessed with a score of 1 (poor), 2 (low), 3 (good), 4 (very good), and 5 (excellent).

Aspect	**Score**				
Pedagogical Knowledge	**5**	**4**	**3**	**2**	**1**
1. I know how to design expected learning outcomes in a clear, precise, concise, and measurable way that aligns with the curriculum, accommodates students' needs, and/or considers the diversity of students (including those with disabilities).					
2. I can design initial/diagnostic assessement to determine students' diversity.					
3. I know how to create scoring and/or various rubrics to increase objectivity in assessing various student abilities.					
Technological Knowledge					
4. I know how to use hardware (for example: laptop, tablet/iPad, laptop camera, internet connection, speakers, etc.) and/or solve technical problems related to the hardware.					
5. I know how to use the software (e.g., download appropriate plug-ins/applications, install programs, create breakout rooms on Zoom, etc.) and/or resolve software-related problems.					
Content Knowledge					
6. I know how to master important concepts from the material that will be taught in class to facilitate diverse students.					
7. I know how to sort essential and non-essential material.					
Technological Content Knowledge					
8. I know how to provide material references and/or a glossary of special terms, which can be accessed online					
Pedagogical Content Knowledge					
9. I know how to start learning so that all students are involved by asking important questions related to the content of the learning material, which will be clarified and answered again at the end of the lesson.					
10. I know how to connect material that students already understand (prior knowledge) with the material to be studied.					
11. I know how to convey the material through contextual and real examples (existing in students' real lives).					
Pedagogical Technological Knowledge					
12. I know how to utilize various chat applications (eg WhatsApp, Instagram, Facebook, Twitter, and the like) to improve communication with students and/or support the implementation of learning activities.					
Technological Pedagogical Content Knowledge					
13. I know how to use a Learning Management System (LMS): Google Classroom, Schoology, Canvas by Instructure, RuangKelas, Edmodo, Moodle, and/or the like to deliver certain material.					
14. I know how to modify assessments for certain materials using an online format (for example: using Google forms, quizzes, Kahoot or other forms of online assessment).					
15. I know how to design alternative project formats: oral presentations, videos, newspaper articles, photo essays, radio documentaries, community research, web publications, etc., for specific materials					
Inclusive Education Knowledge					

16. I know how to instill a sense of tolerance between students in the classroom, including culture and practice.					
17. I know how to manage the class so that all students are comfortable and/or feel welcome in the class.					
18. I know how to respond to students' feelings regarding their learning process.					