Linear: Journal of Mathematics Education
Volume 7, No. 1, June 2026
p-ISSN: 2722-6913 e-ISSN: 2722-760X
https://doi.org/10.32332/linear.v7i1.65-81

Unlocking AI acceptance among pre-service mathematics teachers: development and validation of a TAM-based instrument

Annisa Dwi Kurniawati[1)*], Mega Melinda Ayu Pratiwi[2)]
[1) 2)] Universitas Islam Negeri Kiai Ageng Muhammad Besari Ponorogo, Ponorogo, Indonesia
*annisadwik@uinponorogo.ac.id

Received: 19 January 2026 | Revised: 11 Apryl 2026 | Accepted: 10 May 2026 | Published Online: 06 June 2026

Abstract

The rapid integration of artificial intelligence (AI) in mathematics education has shifted instructional practices beyond efficiency toward more transformative learning experiences. However, the successful adoption of AI depends largely on pre-service teachers' acceptance of these technologies. This study aimed to develop and validate a Technology Acceptance Model (TAM)–based instrument to measure pre-service mathematics teachers' acceptance of AI in transformative learning contexts. A research and development design employing the 4D model (Define, Design, Develop, Disseminate) was implemented. The instrument focused on four core TAM constructs: Perceived Usefulness of AI (PU-AI), Perceived Ease of Use of AI (PEOU-AI), Attitude Toward Use of AI (ATU-AI), and Behavioural Intention to Use AI (BI-AI), with items contextualised for mathematics learning and transformative pedagogical practices. Content validity was evaluated by two expert validators using the Gregory content validity coefficient, and interrater reliability was assessed to ensure consistency across expert evaluations. The results indicated high content validity (V = 0.90) and strong inter-rater agreement (80%), supporting the adequacy and clarity of the developed items. A pilot test involving 47 pre-service mathematics teachers was conducted to examine the clarity and usability of the instrument. Based on expert feedback and pilot testing, a final instrument comprising 20 items was produced, with each item representing a distinct indicator across the four constructs. Based on the results, this instrument is a potentially context-sensitive tool for assessing AI acceptance among pre-service mathematics teachers. These initial findings may inform future research and evaluation efforts, though further psychometric testing, including internal reliability, construct validity, and factor analysis, is recommended before broader application in AI-supported mathematics education.

Keywords: Artificial Intelligence in Education; Instrument Development; Mathematics Teacher Education; Technology Acceptance Model; Transformative Learning.

INTRODUCTION

Despite the growing integration of AI, where its potential to enhance conceptual understanding, problem-solving, and student engagement has been increasingly documented (Canonigo, 2024; Gabriel et al., 2025; Song et al., 2025), a critical gap remains: little is known about how pre-service mathematics teachers perceive and accept these tools. Without a reliable, context-sensitive instrument to measure such acceptance, efforts to prepare future teachers for

AI-supported classrooms lack an empirical foundation (Runge et al., 2025; Walter, 2024). This study addresses that gap directly (Runge et al., 2025; Walter, 2024). Empirical studies in elementary and secondary mathematics classrooms report improvements in performance and problem-solving abilities when AI tools are meaningfully integrated, while also revealing concerns about ethics, privacy, and over-reliance on technology (Canonigo, 2024; Song et al., 2025; Wang et al., 2025). These developments underscore that successful AI integration depends not only on technical quality but crucially on teachers' and students' acceptance, including their trust, perceived value, and alignment with pedagogical visions (Gustilo et al., 2024; Van Den Berg & Papadopoulos, 2024; Wang et al., 2025). Many AI initiatives fail or remain superficial because users do not fully accept AI as part of their practice, even when tools are technically sound (Gustilo et al., 2024; Runge et al., 2025; Zhang et al., 2023).

Recent work emphasises that AI in mathematics education can support transformative learning rather than just efficiency. "Pragmatic AI" has been proposed as a research agenda where AI helps reshape teacher–student interactions, achievement emotions, and learners' sense of competence and control, thereby enabling more meaningful and emotionally supportive mathematics learning experiences (Gabriel et al., 2025). Classroom studies show that AI tutors and tools can mediate rich interactions between pre-service mathematics teachers, instructors, and mathematical content, prompting them to reconsider instructional decision-making, teacher–AI partnerships, and ways of fostering deeper student engagement beyond answer-getting (Biton & Segal, 2025; Song et al., 2025; Yılmaz et al., 2025). Generative AI activities, such as using ChatGPT for problem posing and refinement, can transform pre-service mathematics teachers' technological pedagogical content knowledge (TPACK) and problem-posing practices, encouraging them to rethink mathematical tasks, real-world relevance, and inquiry-based approaches (Biton & Segal, 2025; Song et al., 2025). Such findings position AI as a catalyst for reconfiguring professional identities and beliefs about teaching and learning mathematics, aligning closely with the core concerns of transformative learning.

While the Technology Acceptance Model (TAM) offers a well-established framework for predicting technology adoption through Perceived Usefulness (PU), Perceived Ease of Use (PEOU), Attitude Toward Use (ATU), and Behavioral Intention (BI), its direct application to AI acceptance among pre-service mathematics teachers remains underexplored. This gap raises a pressing question: Is there an instrument that captures AI acceptance in this specific educational context? The absence of such an instrument constitutes the primary problem this study seeks to address (Alejandro et al., 2024; Dahri et al., 2024; Prieto et al., 2020; Runge et al., 2025; Zhang et al., 2023). Across diverse samples of pre-service teachers, in-service teachers, and students, PU and PEOU consistently emerge as primary predictors of BI toward AI tools, generative AI, and AI-based educational applications (Alejandro et al., 2024; Dahri et al., 2024; Li et al., 2024; Runge et al., 2025; Zhang et al., 2023, 2025). Extended TAM models add constructs such as AI self-efficacy, AI literacy, enjoyment, trust, anxiety, and social influence to better capture the complexities of AI integration (Alejandro et al., 2024; Dahri et al., 2024; Guo et al., 2025; Li et al., 2024; Prieto et al., 2020; Runge et al., 2025; Zhang et al., 2025). For example, AI-related teacher training and AI-TPACK increase perceived usefulness and ease of use, which in turn enhance intentions and actual use of AI in future teaching (Runge et al., 2025), while AI awareness shapes PU, PEOU, attitudes, and usage behaviour among K-12 mathematics teachers adopting generative AI (Wang et al., 2025). These results confirm the relevance of TAM for AI in education and underscore the centrality of user perceptions for successful AI adoption.

However, current TAM-based instruments have notable limitations in understanding pre-service mathematics teachers' engagement with AI in transformative learning. Many studies rely on generic TAM scales that treat AI as a broad educational or productivity technology, with items that could apply to any digital tool (Alejandro et al., 2024; Dahri et al., 2024; Li et al., 2024; Prieto et al., 2020; Runge et al., 2025; Zhang et al., 2023, 2025). Even instruments explicitly developed to measure teachers' acceptance of AI or AI-based assessment tend to operationalise PU, PEOU, ATU, BI, self-efficacy, and anxiety at a general level (e.g., "AI in education", "AI-based assessment"), without embedding mathematics-specific content, representations, or disciplinary practices (Guo et al., 2025; Prieto et al., 2020; Zhang et al., 2023). While these instruments provide robust psychometric evidence and useful insights into overall AI acceptance, they do not fully capture how mathematical problem solving, conceptual understanding, or domain-specific pedagogical strategies interact with AI use in ways that might facilitate or hinder transformative learning for future mathematics teachers (Biton & Segal, 2025; Canonigo, 2024; Gabriel et al., 2025; Gurer, 2021; Song et al., 2025; Yılmaz et al., 2025).

Furthermore, existing TAM-based instruments developed for general technology adoption have not been adequately contextualised for the specific demands of AI integration in mathematics education, particularly regarding the unique pedagogical roles and learning experiences of pre-service mathematics teachers. This study, therefore, focuses on developing a TAM-based instrument that is contextually grounded in AI-supported mathematics education, rather than extending TAM with additional theoretical constructs. Studies linking TAM to metacognitive self-regulated learning, intrinsic motivation, or engagement with AI (e.g., ChatGPT as a metacognitive tool) show that AI can support reflective and self-directed learning processes (Dahri et al., 2024; Li et al., 2024). Yet these works still measure acceptance primarily through standard PU, PEOU, ATU, and BI items, occasionally supplemented with enjoyment, trust, or anxiety, rather than operationalising critical reflection, perspective transformation, or shifts in professional identity as mathematics teachers (Biton & Segal, 2025; Dahri et al., 2024; Gabriel et al., 2025; Gurer, 2021; Song et al., 2025). Research on pre-service mathematics teachers' pedagogical beliefs and technology acceptance suggests that constructivist beliefs significantly influence TAM constructs and intentions to use technology (Gurer, 2021), indicating that transformative orientations matter, but these beliefs are not integrated into AI-specific acceptance instruments. Consequently, current measures cannot distinguish between accepting AI as a convenient add-on and accepting AI as a driver of deeper pedagogical change in mathematics education.

This leads to a clear research gap: (a) TAM has been extensively applied to AI in education, but predominantly with general-purpose instruments; (b) there are few instruments targeting AI acceptance specifically in mathematics education, and those that do exist still focus on generic TAM dimensions without systematically embedding the epistemic, affective, and representational particularities of mathematics; and (c) there is virtually no validated TAM-based instrument that explicitly integrates AI, mathematics education, and transformative learning in the context of pre-service mathematics teachers (Biton & Segal, 2025; Canonigo, 2024; Gabriel et al., 2025; Guo et al., 2025; Gurer, 2021; Prieto et al., 2020; Runge et al., 2025; Song et al., 2025; Wang et al., 2025; Yılmaz et al., 2025; Zhang et al., 2023, 2025). Addressing this gap is essential both for advancing theoretical understanding of how transformative learning processes interact with TAM constructs and for providing teacher educators with diagnostic tools to design and evaluate AI-rich, critically reflective mathematics teacher education.

In response, the present study aims to develop a TAM-based instrument specifically contextualised to AI integration in mathematics education, designed to measure pre-service mathematics teachers' perceived usefulness, perceived ease of use, attitude toward use, and behavioral intention regarding AI-supported learning. The novelty of this study lies not in the addition of new theoretical constructs, but in the contextual adaptation of TAM to the specific pedagogical setting of pre-service mathematics teacher education in the context of AI adoption. The instrument retains the core TAM constructs of Perceived Usefulness, Perceived Ease of Use, Attitude Toward Use, and Behavioral Intention, while embedding mathematics-specific and transformative-learning-oriented items that reflect how pre-service mathematics teachers encounter AI as a catalyst for rethinking their instructional practices, beliefs, and professional identities.

Theoretically, this study contributes to the literature by integrating the Technology Acceptance Model with transformative learning perspectives in the specific context of AI-supported mathematics education, thereby extending TAM beyond conventional technology use to encompass deeper transformative learning processes. Methodologically, the study offers a systematic model for instrument development by combining the 4D approach with Gregory-based content validity analysis and inter-rater reliability procedures, which can serve as a reference for future psychometric development in educational research. In practice, the resulting questionnaire serves as a ready-to-use tool for mathematics education programs to evaluate students' acceptance of and perceptions of AI use in their courses, particularly in pre-service mathematics teacher education.

METHOD

This study employed a research-and-development design based on the 4D model (Define, Design, Develop, Disseminate). The 4D model was systematically implemented to guide the stages of needs analysis, instrument construction, expert validation, pilot testing, and dissemination of the final product. Each phase was adapted to meet psychometric standards for educational measurement, particularly for the development of an assessment tool in mathematics education. The research was conducted in the Mathematics Education Study Program at UIN Kiai Ageng Muhammad Besari Ponorogo. The participants in this study were undergraduate students in the mathematics education program. Sampling procedures, inclusion criteria, and sample size were determined to ensure that the participants adequately represented pre-service mathematics teachers in this institutional context. Data analysis procedures were carried out in accordance with the objectives of each development phase. Descriptive statistics were used to examine item distributions, response patterns, and instrument clarity. Reliability analyses were conducted to evaluate the instrument's internal consistency, and additional psychometric analyses were employed where relevant to refine the developed product.

Content validity was established using expert judgment and quantified with the Gregory content validity coefficient. A panel of subject-matter experts was asked to evaluate each item for its relevance and representativeness of the intended construct. Experts classified each item as relevant or irrelevant, and the Gregory index was computed from the proportion of agreement and disagreement among them. Items that did not meet the predetermined content validity criteria were revised or eliminated prior to subsequent testing stages. Inter-rater reliability was examined to ensure consistency among expert evaluations in the content validation process. Agreement between raters was assessed using the percentage agreement method, reflecting the proportion of items on which all raters provided concordant judgements.

RESULTS AND DISCUSSION

Define Stage

Needs Analysis

The needs analysis confirmed that AI tools had already been introduced into mathematics learning activities (e.g., for problem solving, visualisation, feedback, and drill), yet not all pre-service mathematics teachers reported positive acceptance or regular use of these tools. Students indicated both perceived benefits (such as support for conceptual understanding and efficiency in solving tasks) and concerns, including over-reliance on AI, uncertainty about accuracy, and difficulties integrating AI meaningfully into learning processes. These findings are consistent with prior work showing that AI can enhance mathematical learning, but also raise pedagogical and emotional challenges for learners and teachers (Canonigo, 2024; Gabriel et al., 2025; Stefanova & Georgiev, 2024; Yi et al., 2024).

In the local context, no standardised instrument was identified that specifically measured pre-service mathematics teachers' acceptance of AI within a transformative learning orientation (i.e., focusing on conceptual change, critical reflection, and changes in teaching beliefs). Existing TAM-based instruments were developed for general technology adoption and have not been contextualised to the pedagogical demands and disciplinary characteristics of mathematics teacher education (Canonigo, 2024; Li et al., 2024). This instrument-level gap, identified through literature review rather than empirical field data, constitutes the primary justification for the development of a contextualised TAM-based instrument targeted specifically at pre-service mathematics teachers engaging with AI-supported learning.

Construct Identification

Based on the needs analysis and the Technology Acceptance Model, four core constructs were retained as the conceptual basis of the instrument: (a) Perceived Usefulness of AI (PU-AI): the extent to which pre-service mathematics teachers believe that using AI will improve the effectiveness, depth, and quality of mathematics learning, including conceptual understanding, problem solving, and reflective thinking (Canonigo, 2024; Li et al., 2024)(b) Perceived Ease of Use of AI (PEOU-AI): the degree to which AI tools are perceived as easy to learn, operate, and integrate into mathematics learning activities without excessive cognitive or technical effort (Canonigo, 2024; Li et al., 2024)(c) Attitude Toward Use of AI (ATU-AI): overall positive or negative evaluations and feelings about using AI in mathematics learning environments, including openness to experimentation and perceived compatibility with meaningful and transformative learning (Canonigo, 2024; Li et al., 2024)(d) Behavioral Intention to Use AI (BI-AI): the strength of intention to continue using, or to adopt, AI tools in future mathematics teaching and learning practices, particularly for facilitating higher-order and transformative learning experiences (Canonigo, 2024; Li et al., 2024).

Each construct was operationalised to explicitly refer to the use of AI in mathematics learning tasks (e.g., exploring multiple solution strategies, visualising abstract concepts, engaging in reflective dialogue with AI systems)

Indicator Formulation

Indicators for each construct were formulated by synthesising classical TAM indicators, extensions of TAM in AI education, and theoretical work on AI-supported and transformative mathematics learning (Canonigo, 2024; Gabriel et al., 2025; Li et al., 2024; Stefanova & Georgiev, 2024; Yi et al., 2024). For PU-AI, indicators captured perceived enhancement of conceptual understanding, problem-solving performance, engagement, and critical reflection. For PEOU-AI, indicators reflected ease of learning AI tools, clarity of interfaces, and perceived effort required to integrate AI into coursework. ATU-AI indicators reflected enjoyment, alignment of perceived value with good teaching, and willingness to explore AI-based pedagogies. BI-AI indicators described future use intentions and peer recommendations for AI. An indicative summary of the indicator formulation is presented in Table 1.

Table 1. Instrument Indicators

Construct	Indicators	Citations
PU-AI	1. Perceived enhancement of learning engagement and enjoyment through AI-supported mathematics learning 2. Perceived efficiency improvement in completing mathematics tasks and accelerating learning through AI 3. Perceived support for self-directed and autonomous mathematics learning enabled by AI 4. Perceived effectiveness of AI in facilitating error detection, correction, and conceptual understanding in mathematics 5. Perceived usefulness of AI-generated feedback in improving mathematics learning outcomes	(Canonigo, 2024; Gabriel et al., 2025; Li et al., 2024; Stefanova & Georgiev, 2024; Yi et al., 2024)

PEOU-AI	1. Perceived ease of learning to use AI for mathematics learning activities 2. Perceived user-friendliness and operational simplicity of AI interfaces in mathematics learning 3. Perceived independence in using AI for mathematics learning without external support 4. Perceived low cognitive and operational effort required to use AI in mathematics learning 5. Perceived time efficiency in mathematics learning preparation through AI use	(Canonigo, 2024; Gupta et al., 2025; Li et al., 2024)
ATU-AI	1. Positive affective response toward the use of AI in mathematics learning 2. Positive evaluative belief regarding the impact of AI on mathematics learning 3. Favourable disposition toward the future adoption of AI in mathematics learning 4. Positive emotional evaluation of AI use in enhancing enjoyment in mathematics learning 5. Positive overall attitude toward AI as a facilitator of success in mathematics learning	(Canonigo, 2024; Deng et al., 2019, p. 20; Li et al., 2024)
BI-AI	1. Intention to recommend and promote the use of AI for mathematics learning to others 2. Intention to integrate AI into the preparation and implementation of mathematics teaching and learning activities 3. Willingness to explore and adopt new AI features for mathematics learning purposes 4. Intention to invest time and effort in developing competencies related to AI use in learning contexts 5. Long-term commitment to continuous learning and professional development related to AI in education	(Canonigo, 2024; Li et al., 2024)

Design Stage

Instrument Blueprint

A detailed instrument blueprint (table of specifications) was produced linking constructs, indicators, and item codes. Each indicator within each construct was initially represented by two to three items in the instrument. The initial allocation yielded a balanced distribution of items across constructs to ensure adequate coverage and redundancy for subsequent psychometric screening, following best-practice recommendations for scale development (Gupta et al., 2025; Stefana et al., 2025; Zickar, 2020). An illustrative specification is presented in Table 2.

Table 2. Instrument Blueprint

Construct	Indicators	Initial Items
PU-AI	1. Perceived enhancement of learning engagement and enjoyment through AI-supported mathematics learning	10

Construct	Indicators	Initial Items
	2. Perceived efficiency improvement in completing mathematics tasks and accelerating learning through AI 3. Perceived support for self-directed and autonomous mathematics learning enabled by AI 4. Perceived effectiveness of AI in facilitating error detection, correction, and conceptual understanding in mathematics 5. Perceived usefulness of AI-generated feedback in improving mathematics learning outcomes	
PEOU-AI	1. Perceived ease of learning to use AI for mathematics learning activities 2. Perceived user-friendliness and operational simplicity of AI interfaces in mathematics learning 3. Perceived independence in using AI for mathematics learning without external support 4. Perceived low cognitive and operational effort required to use AI in mathematics learning 5. Perceived time efficiency in mathematics learning preparation through AI use	15
ATU-AI	1. Positive affective response toward the use of AI in mathematics learning 2. Positive evaluative belief regarding the impact of AI on mathematics learning 3. Favourable disposition toward the future adoption of AI in mathematics learning 4. Positive emotional evaluation of AI use in enhancing enjoyment in mathematics learning 5. Positive overall attitude toward AI as a facilitator of success in mathematics learning	15
BI-AI	1. Intention to recommend and promote the use of AI for mathematics learning to others 2. Intention to integrate AI into the preparation and implementation of mathematics teaching and learning activities 3. Willingness to explore and adopt new AI features for mathematics learning purposes 4. Intention to invest time and effort in developing competencies related to AI use in learning contexts 5. Long-term commitment to continuous learning and professional development related to AI in education	10

Initial Item Pool

The initial item pool consisted of approximately 50 Likert-type statements, evenly covering the four constructs. Items were written in clear, student-appropriate English and embedded in mathematics learning scenarios involving AI (e.g., using AI to explore multiple solution methods, to receive hints, or to visualise graphs). Item wording followed principles of clarity, simplicity, and unidimensionality as recommended in measurement guidelines (Gupta et al., 2025; Stefana et al., 2025; Zickar, 2020). Examples of PU-AI items included statements such as "AI helps me complete mathematics tasks or exercises and learn mathematics more quickly."

Format and Scale

All items were formatted as closed-ended statements rated on a five-point Likert scale ranging from 1 ("strongly disagree") to 5 ("strongly agree"). The questionnaire began with brief instructions and demographic questions (e.g., semester, prior AI experience), followed by items grouped by the four TAM constructs. This format aligns with standard practice in technology acceptance and educational scale development (Canonigo, 2024; Deng et al., 2019, p. 20; Li et al., 2024).

Develop Stage

Content Validity (Gregory Analysis)

Content validity was evaluated by a panel of two experts in mathematics education and educational technology/AI, consistent with recommendations to use expert panels for content validation (Gupta et al., 2025; Pueyo-Garrigues et al., 2020; Stefana et al., 2025). Each expert rated the item's relevance to its intended construct (relevant/irrelevant). Using the Gregory coefficient, indices were calculated for each item.

Most items achieved Gregory indices in the acceptable to high range (e.g., V ≥ 0.75), comparable to values reported for other educational and evaluation instruments (Qin et al., 2025; Setiawan et al., 2024). A summary table for the content validity is presented in Table 3.

Table 3. Summary of Content Validity Gregory

Expert Judgement Results		**Rater I**	
		Less Relevant (Score 1-2)	**Highly Relevant (Score 3-4)**
Rater II	Less Relevant (Score 1-2)	0	0
	Highly Relevant (Score 3-4)	1	9

Content validation was conducted by two expert validators. The results of the expert judgment are presented in Table 3. Based on the Gregory formula, the content validity coefficient was calculated as follows.

$$V = \frac{9}{0 + 0 + 1 + 9} = 0{,}9$$

The obtained content validity index of 0.90 indicates a high level of agreement between the two validators. Therefore, the developed questionnaire is content valid and suitable for use.

Furthermore, based on the expert validators' suggestions and feedback, each indicator within each construct was adequately represented by a single item. As a result, the final version of the instrument comprised 20 items, each representing a distinct indicator, as presented in Table 4.

Table 4. Final Items of the Instrument

Construct	Statement Item	Final Items
PU-AI	1. Using AI makes mathematics learning more engaging and enjoyable 2. AI helps me complete mathematics tasks and learn mathematics more efficiently 3. AI supports me in learning mathematics independently without relying heavily on others 4. AI helps me identify and correct mistakes, as well as deepen my understanding of mathematical concepts 5. Using AI provides useful feedback that helps me achieve better mathematics learning outcomes	5
PEOU-AI	1. I find AI technology easy to learn for use in mathematics learning 2. I find the AI interface in mathematics learning user-friendly and easy to operate 3. I can use AI in mathematics learning without much assistance from others 4. I feel that using AI in mathematics learning does not require much additional effort 5. I feel that using AI helps save time in preparing mathematics learning materials	5
ATU-AI	1. I feel pleased, satisfied, and enthusiastic when using AI in mathematics learning 2. I believe that AI has a positive impact on mathematics learning 3. I support the application of AI in mathematics learning in the future 4. I believe that using AI makes mathematics learning more enjoyable 5. I feel that using AI supports my success in mathematics learning	5
BI-AI	1. I would recommend the use of AI for mathematics learning to my peers or colleagues 2. I plan to use AI in preparing and designing mathematics learning materials and integrating it with instructional methods 3. I am willing to explore new features offered by AI to support mathematics learning 4. I am willing to dedicate time to learning more about the use of AI in education 5. I am committed to continuously updating my knowledge of AI to support mathematics learning	5

Inter-Rater Reliability

The instrument's reliability in this study was established using an inter-rater agreement (IRA) approach, involving two independent raters. IRA quantifies the extent to which different raters assign the same score to the same items, and is widely used to evaluate the consistency and trustworthiness of observational or expert-based ratings in education, psychology, and health research (McHugh, 2012; O'Neill, 2017; Shweta et al., 2015). In line with common practice, percent agreement was used as the index of IRA by counting the number of items scored identically by both raters and dividing this by the total number of items, then expressing the result as a percentage (McHugh, 2012; Shweta et al., 2015; Westergård et al., 2019). The reliability estimate based on inter-rater agreement was calculated using the following formula.

$$Inter - rater\ agreement = \frac{Number\ of\ items\ scored\ identically\ by\ both\ raters}{Total\ number\ of\ items} \times 100\%$$

After the content validation phase, the questionnaire was subjected to this reliability test to ensure consistency of judgments between raters. Two raters with expertise in mathematics education independently scored the instrument items. Based on their ratings, the inter-rater agreement was:

$$Inter - rater\ agreement = \frac{8}{10} \times 100\% = 80\%$$

An IRA value of 80% indicates a high level of exact agreement between the two raters on the questionnaire items. Percent agreement around or above 0.80 is often cited as an acceptable benchmark for inter-rater agreement in applied educational and observational research, particularly when percent agreement or "percent within one" is used as the main reliability index (McHugh, 2012; Shweta et al., 2015; Westergård et al., 2019; Wilhelm et al., 2018). In methodological reviews, percent agreement is described as an intuitive and interpretable index of agreement, especially when the goal is to ensure that raters produce the same absolute scores, even though it does not correct for chance agreement (McHugh, 2012; O'Neill, 2017; Shweta et al., 2015; Zhao et al., 2022).

Given that values of IRA ≥ 0.75 are commonly interpreted as indicating good agreement and values ≥ 0.90 as very high or "almost perfect" agreement in many applied contexts (Abbott et al., 2024; McHugh, 2012; Shweta et al., 2015; Westergård et al., 2019). The obtained IRA of 80% supports the conclusion that the developed questionnaire has high inter-rater reliability/agreement and is suitable for data collection in terms of both content consistency and rating objectivity (Abbott et al., 2024; McHugh, 2012; O'Neill, 2017; Shweta et al., 2015; Wilhelm et al., 2018).

Pilot Test Results

A pilot study was conducted with approximately 47 pre-service mathematics teachers. Descriptive analysis showed that all response options were used and that item distributions were acceptable (no extreme ceiling/floor effects).

Final Instrument

After iterative revisions based on content validity, interrater reliability, and pilot testing, a final instrument was developed. The final version contained a reduced and optimised number of items distributed across the four TAM constructs. Each construct was represented by five items, each reflecting the construct's indicators.

Disseminate Stage

Final Version of the Instrument

The final questionnaire was formatted as a structured instrument with clear instructions, construct labels, and item codes (e.g., PU1–PU5, PEOU6–PEOU10, ATU11–ATU15, BI16–B20). Items were grouped by construct for ease of administration and scoring. A scoring guide specifying how to compute subscale means and interpret higher scores on each construct was prepared, following conventions in validated scales (Deng et al., 2019; Gümüş & Kukul, 2022; Joly et al., 2025; Qin et al., 2025; Wongpakaran et al., 2020).

Instrument Readiness

Based on the accumulated evidence of strong content validity and satisfactory inter-rater reliability, the instrument was deemed psychometrically sound and ready for broader application. Reliability and validity indices were comparable to, or exceeded, benchmarks reported for similar instruments in digital competence, AI-related pedagogy, and educational evaluation (Alqarni, 2025; Deng et al., 2019; Gümüş & Kukul, 2022; Joly et al., 2025; Qin et al., 2025; Setiawan et al., 2024; Wongpakaran et al., 2020).

Potential Applications

The validated instrument is suitable for use in multiple contexts of mathematics education and teacher preparation, including: (a) Research on pre-service mathematics teachers' acceptance of AI and its relationship to learning outcomes or teaching beliefs (Canonigo, 2024; Gabriel et al., 2025; Stefanova & Georgiev, 2024; Yi et al., 2024; Yılmaz et al., 2025); (b) Evaluation of AI-integrated courses, teaching innovations, and professional development programs in mathematics education (Deng et al., 2019; Gümüş & Kukul, 2022; Qin et al., 2025;

Setiawan et al., 2024; Xu, 2024); (c) Needs assessment and policy development related to AI adoption in teacher education curricula and institutional strategies for AI integration (Gabriel et al., 2025; Stefanova & Georgiev, 2024; Xu, 2024; Yi et al., 2024; Yılmaz et al., 2025).

These findings carry significant implications for mathematics education, where AI acceptance cannot be treated as a generic technology adoption phenomenon. The unique nature of mathematics as a discipline (characterised by its abstract reasoning demands, procedural complexity, and historically high levels of student anxiety) necessitates an instrument that captures AI acceptance within a subject-specific pedagogical context, rather than relying on generalised technology acceptance measures. This contextual specificity is precisely what distinguishes the present instrument from existing TAM-based tools developed for broader educational technology settings. These potential applications align with broader calls for robust, context-sensitive instruments to monitor AI integration in education and inform evidence-based practice and policy (Canonigo, 2024; Gabriel et al., 2025; Stefanova & Georgiev, 2024; Xu, 2024; Yi et al., 2024; Yılmaz et al., 2025).

CONCLUSION

This study concludes that a Technology Acceptance Model (TAM) based instrument was successfully developed to assess pre-service mathematics teachers' acceptance of artificial intelligence in transformative mathematics learning using the 4D development model. The validation results demonstrated strong content validity and high inter-rater reliability, indicating that the instrument reliably represents essential dimensions of AI acceptance, including perceived usefulness, perceived ease of use, attitudes, and behavioural intentions. The primary contribution of this research is the development of a context-specific, empirically validated measurement tool that addresses the growing need for a systematic evaluation of AI integration in mathematics teacher education. This instrument provides a rigorous foundation for future empirical investigations and institutional decision-making related to AI-supported learning design. Further research is recommended to implement the instrument across broader institutional and cultural contexts and to examine its association with pedagogical transformation and student learning outcomes. It is expected that the findings of this study will support evidence-based implementation of artificial intelligence in mathematics education, promoting innovative, reflective, and sustainable teaching practices.

REFERENCES

Abbott, A., Shin, J., Carlson, K., Russell, M., Qi, Y., Storm, H., & Jewell, V. (2024). Achieving

Inter-Rater Agreement and Inter-Rater Reliability to Assess Fidelity of an Occupation-Based Coaching (OBC) Clinical Trial Intervention. *British Journal of Occupational Therapy*, *88*, 133–141. https://doi.org/10.1177/03080226241283292

Alejandro, I. M., Sánchez, J. M., Sumalinog, G., Mananay, J., Goles, C., & Fernández, C. (2024). Pre-service teachers' technology acceptance of artificial intelligence (AI) applications in education. *STEM Education*. https://doi.org/10.3934/steme.2024024

Alqarni, A. (2025). Artificial Intelligence-Critical Pedagogic: Design and Psychologic Validation of a Teacher-Specific Scale for Enhancing Critical Thinking in Classrooms. *J. Comput. Assist. Learn.*, *41*. https://doi.org/10.1111/jcal.70039

Biton, Y., & Segal, R. (2025). Learning to Craft and Critically Evaluate Prompts: The Role of Generative AI (ChatGPT) in Enhancing Pre-service Mathematics Teachers' TPACK and Problem-Posing Skills. *International Journal of Education in Mathematics, Science and Technology*. https://doi.org/10.46328/ijemst.4654

Canonigo, A. M. (2024). Levering AI to enhance students' conceptual understanding and confidence in mathematics. *Journal of Computer Assisted Learning*, *40*(6), 3215–3229. https://doi.org/10.1111/jcal.13065

Dahri, N. A., Yahaya, N., Al-Rahmi, W., Aldraiweesh, A., Alturki, U., Almutairy, S., Shutaleva, A., & Soomro, R. B. (2024). Extended TAM based acceptance of AI-Powered ChatGPT for supporting metacognitive self-regulated learning in education: A mixed-methods study. *Heliyon*, *10*. https://doi.org/10.1016/j.heliyon.2024.e29317

Deng, R., Benckendorff, P., & Gannaway, D. (2019). Learner engagement in MOOCs: Scale development and validation. *Br. J. Educ. Technol.*, *51*, 245–262. https://doi.org/10.1111/bjet.12810

Gabriel, F., Kennedy, J., Marrone, R., & Leonard, S. (2025). Pragmatic AI in education and its role in mathematics learning and teaching. *NPJ Science of Learning*, *10*. https://doi.org/10.1038/s41539-025-00315-4

Gümüş, M. M., & Kukul, V. (2022). Developing a digital competence scale for teachers: Validity and reliability study. *Education and Information Technologies*, *28*, 2747–2765. https://doi.org/10.1007/s10639-022-11213-2

Guo, S., Shi, L., & Zhai, X. (2025). Developing and validating an instrument for teachers' acceptance of artificial intelligence in education. *Education and Information Technologies*, *30*, 13439–13461. https://doi.org/10.1007/s10639-025-13338-6

Gupta, S., Grover, S., Menon, V., Indu, P., Chacko, D., & S, M. G. (2025). Item generation and establishing face and content validity of a rating scale: A primer. *Indian Journal of Psychiatry*, *67*, 816–822. https://doi.org/10.4103/indianjpsychiatry_750_25

Gurer, M. D. (2021). Examining technology acceptance of pre-service mathematics teachers in Turkey: A structural equation modeling approach. *Education and Information Technologies*, *26*, 4709–4729. https://doi.org/10.1007/s10639-021-10493-4

Gustilo, L., Ong, E., & Lapinid, M. (2024). Algorithmically-driven writing and academic integrity: Exploring educators' practices, perceptions, and policies in AI era. *International Journal for Educational Integrity*, *20*, 1–43. https://doi.org/10.1007/s40979-024-00153-8

Joly, B., Gray, C., Cousineau, K., Pearson, K., & Harder, V. (2025). The development, validity, and reliability of the Researcher Investment Tool. *Journal of Clinical and Translational Science*, *9*. https://doi.org/10.1017/cts.2024.673

Li, W., Zhang, X., Li, J., Yang, X., Li, D., & Liu, Y. (2024). An explanatory study of factors influencing engagement in AI education at the K-12 Level: An extension of the classic TAM model. *Scientific Reports*, *14*. https://doi.org/10.1038/s41598-024-64363-3

McHugh, M. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*, 276–282. https://doi.org/10.11613/bm.2012.031

O'Neill, T. (2017). An Overview of Interrater Agreement on Likert Scales for Researchers and Practitioners. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00777

Prieto, J. C. S., Cruz-Benito, J., Therón, R., & García-Peñalvo, F. (2020). Assessed by Machines: Development of a TAM-Based Tool to Measure AI-based Assessment Acceptance Among Students. *Int. J. Interact. Multim. Artif. Intell.*, *6*, 80–86. https://doi.org/10.9781/ijimai.2020.11.009

Pueyo-Garrigues, M., Pardavila-Belio, M., Whitehead, D., Esandi, N., Canga-Armayor, A., Elosua, P., & Canga-Armayor, N. (2020). Nurses' knowledge, skills and personal attributes for competent health education practice: An instrument development and psychometric validation study. *Journal of Advanced Nursing*. https://doi.org/10.1111/jan.14632

Qin, L., Ho, W. K. Y., & Khoo, S. (2025). Development and validation of a measurement instrument for student assessment of quality physical education in Chinese secondary schools. *PLOS One*, *20*. https://doi.org/10.1371/journal.pone.0324227

Runge, I., Hebibi, F., & Lazarides, R. (2025). Acceptance of Pre-Service Teachers Towards Artificial Intelligence (AI): The Role of AI-Related Teacher Training Courses and AI-TPACK Within the Technology Acceptance Model. *Education Sciences*. https://doi.org/10.3390/educsci15020167

Setiawan, R., Wagiran, W., & Alsamiri, Y. (2024). Construction of an instrument for evaluating the teaching process in higher education: Content and construct validity. *REID (Research and Evaluation in Education)*. https://doi.org/10.21831/reid.v10i1.63483

Shweta, Bajpai, R., & Chaturvedi, H. (2015). Evaluation of Inter-Rater Agreement and Inter-Rater Reliability for Observational Data: An Overview of Concepts and Methods. *Journal of the Indian Academy of Applied Psychology*, *41*, 20.

Song, X., Mak, J., & Chen, H. (2025). Teachers and Learners' Perceptions about Implementation of AI Tools in Elementary Mathematics Classes. *SAGE Open*, *15*(2), 21582440251334545. https://doi.org/10.1177/21582440251334545

Stefana, A., Damiani, S., Granziol, U., Provenzani, U., Solmi, M., Youngstrom, E., & Fusar-Poli, P. (2025). Psychological, psychiatric, and behavioral sciences measurement scales: Best practice guidelines for their development and validation. *Frontiers in Psychology*, *15*. https://doi.org/10.3389/fpsyg.2024.1494261

Stefanova, T., & Georgiev, S. (2024). Possibilities for using AI in mathematics education. *Mathematics and Education in Mathematics*. https://doi.org/10.55630/mem.2024.53.117-125

Van Den Berg, S., & Papadopoulos, P. (2024). Summative assessment with artificial intelligence: Qualitative analysis and comparison of technology acceptance in student and teacher populations. *Innovations in Education and Teaching International*, *62*, 1529–1544. https://doi.org/10.1080/14703297.2024.2436613

Walter, Y. (2024). Embracing the future of Artificial Intelligence in the classroom: The relevance of AI literacy, prompt engineering, and critical thinking in modern education. *International Journal of Educational Technology in Higher Education*, *21*, 1–29. https://doi.org/10.1186/s41239-024-00448-3

Wang, Y., Wei, Z., Wijaya, T. T., Cao, Y., & Ning, Y. (2025). Awareness, acceptance, and adoption of Gen-AI by K-12 mathematics teachers: An empirical study integrating TAM and TPB. *BMC Psychology*, *13*. https://doi.org/10.1186/s40359-025-02781-2

Westergård, E., Ertesvåg, S., & Rafaelsen, F. (2019). A Preliminary Validity of the Classroom Assessment Scoring System in Norwegian Lower-Secondary Schools. *Scandinavian Journal of Educational Research*, *63*, 566–584. https://doi.org/10.1080/00313831.2017.1415964

Wilhelm, A., Rouse, A., & Jones, F. (2018). Exploring Differences in Measurement and Reporting of Classroom Observation Inter-Rater Reliability. *Practical Assessment, Research and Evaluation*, *23*, 4.

Wongpakaran, N., Wongpakaran, T., Pinyopornpanish, M., Simcharoen, S., Suradom, C., Varnado, P., & Kuntawong, P. (2020). Development and validation of a 6-item Revised UCLA Loneliness Scale (RULS-6) using Rasch analysis. *British Journal of Health Psychology*. https://doi.org/10.1111/bjhp.12404

Xu, Z. (2024). AI in education: Enhancing learning experiences and student outcomes. *Applied and Computational Engineering*. https://doi.org/10.54254/2755-2721/51/20241187

Yi, L., Liu, D., Jiang, T., & Xian, Y. (2024). The Effectiveness of AI on K-12 Students' Mathematics Learning: A Systematic Review and Meta-Analysis. *International Journal of Science and Mathematics Education*, *23*, 1105–1126. https://doi.org/10.1007/s10763-024-10499-7

Yılmaz, Z., Galanti, T., Naresh, N., & Kanbir, S. (2025). Exploring the interactions among instructor, prospective teachers and AI in facilitating mathematics learning. *School Science and Mathematics*. https://doi.org/10.1111/ssm.18341

Zhang, C., Hu, M., Wu, W., Chen, Y., Kamran, F., & Wang, X. (2025). A profile analysis of pre-service teachers' AI acceptance: Combining behavioral, technological, and human factors. *Teaching and Teacher Education*. https://doi.org/10.1016/j.tate.2025.105086

Zhang, C., Schießl, J., Plößl, L., Hofmann, F., & Gläser-Zikuda, M. (2023). Acceptance of artificial intelligence among pre-service teachers: A multigroup analysis. *International Journal of Educational Technology in Higher Education*, *20*. https://doi.org/10.1186/s41239-023-00420-7

Zhao, X., Feng, G., Ao, S., & Liu, P. L. (2022). Interrater reliability estimators tested against true interrater reliabilities. *BMC Medical Research Methodology*, *22*. https://doi.org/10.1186/s12874-022-01707-5

Zickar, M. (2020). Measurement Development and Evaluation. *Annual Review of Organizational Psychology and Organizational Behavior*. https://doi.org/10.1146/annurev-orgpsych-012119-044957