Contents lists available at https://ojs.unsulbar.ac.id/index.php/saqbe
Vol 2 No 1 Maret 2025, https://doi.org/10.31605/saqbe

Saqbe : Sains dan Pembelajarannya

Validation of Pre-Service STEM Teachers' Acceptance and Use of Generative Artificial Intelligence Scale: Rasch Model

Validasi Skala Penerimaan dan Pemanfaatan Kecerdasan Buatan Generatif oleh Calon Guru STEM: Analisis Model Rasch

Elsima Nainggolan[1], Saraswathy A/p Ramasundrum[2], Indra Maulana[3]
[1]Chemistry Education, Faculty of Sciences, Universitas Negeri Yogyakarta, Indonesia
[2]Science Education, Faculty of Education, Universiti Malaya, Malaysia
[3]Informatics and Computer Engineering, Institut Prima Bangsa, Indonesia
*Correspondence e-mail: elsimanainggolan.2022@student.uny.ac.id

Received: 1 January 2025
Revised: 30 March 2025
Accepted : 30 March 2025
Published: 30 March 2025

Abstrak

Mahasiswa calon guru STEM dari Generasi Z, yang dikenal sebagai generasi native digital, memiliki potensi kuat untuk mengadopsi Kecerdasan Buatan Generatif (GAI) dalam konteks pendidikan dan mendorong integrasinya secara bermakna. Sehubungan dengan hal tersebut, penelitian ini bertujuan untuk memvalidasi sebuah instrumen yang dirancang untuk mengukur penerimaan dan penggunaan GAI. Penelitian ini menggunakan survei kuantitatif cross-sectional yang melibatkan 401 mahasiswa calon guru STEM dengan instrumen berbasis UTAUT2–TPB. Data dikumpulkan melalui Google Forms dan dianalisis menggunakan pemodelan Rasch (Winsteps 3.7.3). Hasil penelitian mengonfirmasi bahwa instrumen memiliki karakteristik psikometrik yang kuat berdasarkan model Rasch. Analisis menunjukkan reliabilitas person (0,94) dan item (0,97) yang tinggi, indeks pemisahan yang jelas, serta nilai kesesuaian item yang dapat diterima. Hal ini mengindikasikan bahwa skala mampu membedakan responden secara efektif dan mempertahankan stabilitas antaritem. Uji unidimensionalitas juga menunjukkan bahwa instrumen mengukur satu konstruk yang koheren, sehingga memperkuat integritas struktural internalnya. Secara keseluruhan, temuan ini memverifikasi bahwa instrumen tersebut valid dan reliabel untuk menilai penerimaan dan penggunaan GAI di kalangan mahasiswa calon guru STEM. Penelitian ini memberikan kontribusi teoretis dan praktis dengan menyediakan alat ukur yang teruji secara ketat untuk mengevaluasi kesiapan adopsi dan integrasi GAI dalam pendidikan calon guru.

Kata Kunci

Calon Guru STEM, UTAUT2, TPB, Generative Artificial Intelligence, Rasch Model

Abstract

Generation Z pre-service STEM teachers, recognized as digital natives, possess strong potential to adopt Generative Artificial Intelligence (GAI) in educational contexts and advance its meaningful integration. Accordingly, this study aims to validate an instrument designed to measure their acceptance of and use of GAI. A quantitative cross-sectional survey was administered to 401 pre-service STEM teachers using a UTAUT2–TPB–based instrument. Data were collected via Google Forms and analyzed using Rasch modeling (Winsteps 3.7.3). The results confirm that the instrument possesses strong psychometric properties under the Rasch model. Analysis demonstrated high person (0.94) and item (0.97) reliabilities, well-defined separation indices, and acceptable item fit values, indicating that the scale effectively differentiates respondents and maintains stability across items. The unidimensionality test further supported that the instrument measures a single, coherent construct, reinforcing its internal structural integrity. Overall, these findings verify that the instrument is valid and reliable for assessing GAI acceptance and use among pre-service STEM teachers. The study offers both theoretical and practical contributions by providing a rigorously tested measurement tool to evaluate readiness for GAI adoption and integration in teacher education.

Keywords
Pre-service STEM teachers, UTAUT2, TPB, Generative Artificial Intelligence, Rasch Model

INTRODUCTION

Artificial Intelligence (AI) has rapidly evolved into a transformative force across multiple sectors, including healthcare, finance, and communication, with education increasingly recognized as a domain of profound potential impact (Dewi, Qudratuddarsi, Ningthias, & Cinthami, 2024; Yang, Zhang, Sun, He & Wei, 2025). Beyond the automation of routine tasks, AI has the capacity to reshape how knowledge is delivered, accessed, and assessed. In contemporary classrooms, AI is no longer a futuristic possibility but a tangible and growing reality. Prominent examples include generative platforms such as ChatGPT, adaptive learning environments, and automated grading systems (Wang & Huang, 2025). These technologies facilitate personalized learning, provide immediate feedback, and offer teachers rich learning analytics capable of identifying struggling learners early. Intelligent tutoring systems function as digital teaching assistants, while AI-powered simulations enhance engagement with complex or abstract concepts. For teachers, AI tools also support lesson planning, content development, and differentiated instruction (Yang, Sun, Sun & Salas-Pilco, 2025).

Despite these opportunities, challenges remain. Over-reliance on AI may undermine critical thinking, while algorithmic bias and data privacy raise ethical concerns (Chen & Lin, 2024). Effective adoption also depends on educators' readiness and digital competence. Without adequate preparation, teachers may misapply or underutilize AI tools, limiting their pedagogical value (Schiff, 2022). STEM teachers are particularly strategic in this context, as STEM fields have historically led technological innovation and already rely on simulations, modeling, and digital laboratories (Xu & Ouyang, 2022; Zhai, Neumann & Krajcik, 2023). Their early adopter tendencies, direct disciplinary benefits from AI, and centrality to national innovation agendas position them as key actors in advancing AI integration (Chng, Tan & Tan, 2023; Ouyang, Dinh & Xu, 2023; Parviz, 2024).

Equipping pre-service teachers with AI-related competencies is therefore essential. Their acceptance, attitudes, and intentions toward AI will ultimately influence how AI is used in schools in the coming years. Teacher attitudes strongly shape technology adoption: positive perceptions motivate experimentation, while negative attitudes result in rejection or minimal usage (Özden, Yaşar & Meydan, 2025; Sun, Tian, Sun, Fan & Yang, 2024; Zhang, Schießl, Plößl, Hofmann & Gläser-Zikuda, 2023). For pre-service STEM teachers, readiness is even more consequential, as AI can enhance visualization, simulation, and inquiry-based learning. However, research on teachers remains limited compared to research on students, and studies focusing specifically on pre-service STEM teachers are even scarcer (Guan, Zhang & Gu, 2025; Laru, Celik, Jokela & Mäkitalo, 2025).

Investigating pre-service teachers' relationship with AI requires valid and reliable instruments. Constructs such as "attitude" and "intention" are latent and must be inferred through measurement tools. Poorly validated instruments risk measurement error and conceptual ambiguity, while strong measurement tools support comparability and policy relevance. Established models such as the Theory of Planned Behavior (TPB) and the Unified Theory of Acceptance and Use of Technology 2 (UTAUT2) offer robust conceptual foundations for designing such instruments.Despite growing interest in AI adoption, methodological gaps remain (Hidayat, Qudratuddarsi, Ayub, & Latif, 2025; Saddia, Yanti, & Qudratuddarsi, 2025). Prior studies commonly relied on Classical Test Theory (CTT), drawing on metrics such as Cronbach's alpha. Although useful, CTT has limitations in assessing item-level misfit, ensuring interval scaling, and confirming internal measurement coherence. In contrast, modern psychometric approaches such as the Rasch model offer item-level precision, allow examination of reliability and separation, test unidimensionality, and evaluate measurement fairness through differential item functioning (Hope, Kluth, Homer, Dewar, Goddard-Fuller, Jaap, & Cameron, 2025; Medvedev, & Krägeloh, 2025). However, few studies have applied Rasch modeling to instruments measuring AI adoption in teacher education. To address these gaps, this study validates a new instrument specifically designed to measure pre-service STEM teachers' acceptance and intended use of AI, employing the Rasch model to ensure rigorous psychometric evaluation grounded in modern measurement theory.

METHOD

Research Design

This research adopted a quantitative survey approach, emphasizing the collection and interpretation of numerical data obtained from participants' structured responses. As a survey-oriented study, it aimed to assess pre-service STEM teachers' acceptance and Use Toward Artificial Intelligence at a specific point in time, without introducing external interventions or manipulating participant conditions (Jamieson, Govaart, & Pownall, 2023; Plonsky, 2017). To achieve this, the researchers employed a cross-sectional design, which provided a single-time "snapshot" of participants' skills. This design was particularly advantageous because it minimized complications commonly associated with longitudinal research, such as participant dropout or external contextual changes that may influence outcomes. The application of a quantitative framework ensured a high degree of objectivity, as data were treated in measurable terms, reducing the likelihood of researcher bias (Kesmodel, 2018).

Research Subject

A total of 401 Generation Z pre-service STEM teachers participated in this study, recruited through convenience sampling. While this method offered efficiency and accessibility in reaching participants, it may somewhat limit the generalizability of the results. Nevertheless, the sample was considered highly relevant, as these individuals are accustomed to technology in both their daily lives and academic experiences. Table 1 summarizes the demographic characteristics of the participants, showing that 30.17% were male and 69.83% were female, with the majority being female. Regarding year of study, 28.18% were in their first year, 26.18% in their second year, 32.67% in their third year, and 12.97% in their fourth year.

Table 1. Sample of the study

Sample	N	Percentage
Gender		
Male	121	30.17%
Female	280	69.83%
Year of study		
First year	113	28.18%
Second year	105	26.18%
Third Year	131	32.67%
Fourth year	52	12.97%
Total	401	100 %

Instrument

The instrument used in this study was adapted from previous research (Habibi, Muhaimin, Danibao, Wibowo, Wahyuni & Octavia, 2023; Habibi, Mukminin, Octavia, Wahyuni, Danibao, & Wibowo, 2024) and subsequently revalidated by the researcher to ensure its suitability within the present context. To establish content validity, the instrument was reviewed by two educational technology experts who assessed its appropriateness in measuring the intended constructs and achieving the research objectives. The instrument measuring acceptance and use of Generative Artificial Intelligence was developed by integrating two theoretical frameworks: the Unified Theory of Acceptance and Use of Technology 2 (UTAUT2) and the Theory of Planned Behavior (TPB). The instrument consists of several sub-constructs, namely: Subjective Norms (SN), Performance Expectancy (PE), Effort Expectancy (EE), Hedonic Motivation (HM), Facilitating Conditions (FC), Habit (H), Attitude (AT), Perceived Behavioral Control (PBC), Behavioral Intention (BI), and AI Use (AIU). Each sub-construct was operationalized through multiple items: three items for PE, EE, HM, PBC, BI, and AIU; four items for FC, H, and AT; and five items for SN. This systematic adaptation, validation, and theoretical grounding ensures that the instrument is both contextually relevant and psychometrically robust, providing a reliable tool for measuring pre-service teachers' acceptance and use of Artificial Intelligence.

Data Collection

Data collection was carried out using Google Forms, which aligns with sustainable and environmentally friendly research practices by reducing paper use, while also enhancing efficiency and minimizing potential errors from manual data entry. The use of a digital platform provided the added benefit of real-time response tracking, enabling the researcher to monitor data quality promptly and address any inconsistencies. To further strengthen the reliability of the responses, the researcher maintained an active presence during data collection. This ensured that participants could seek clarification whenever survey items were unclear, thereby reducing misinterpretation and enhancing the construct validity of the instrument. In addition, the supportive presence of the researcher fostered a more comfortable and trustworthy atmosphere, which likely encouraged participants to answer truthfully and thoughtfully. Participation was strictly voluntary, and participants were explicitly informed that their responses would remain confidential and would not affect their academic standing or evaluations in any way. These ethical assurances are consistent with best practices in educational research, as they protect participant autonomy, minimize coercion, and uphold the integrity and credibility of the dataset. By combining digital efficiency, methodological rigor, and strong ethical safeguards, the study ensured both data quality and research trustworthiness.

Data Analysis

The data in this study were analyzed using two complementary approaches: the Rasch model and Confirmatory Factor Analysis (CFA). The Rasch model analysis was conducted with Winsteps version 3.7.3, while CFA was performed using AMOS version 24. Employing both methods allows for a rigorous examination of the psychometric properties of the instrument, addressing both item-level functioning and construct-level validation. In the Rasch analysis, several key aspects were evaluated, including reliability and separation indices, item fit statistics (infit and outfit mean square [MNSQ], point-measure correlation [Pt. Mea. Corr.]), and unidimensionality. Reliability and separation indices provide evidence of the instrument's capacity to consistently differentiate between respondents and to establish a meaningful hierarchy of items. Item fit statistics ensure that individual items align with the expectations of the Rasch measurement model, thereby supporting construct validity. Unidimensionality testing is critical, as it verifies that the items collectively measure a single underlying latent construct, which is a fundamental assumption of Rasch modeling.

RESULT AND DISCUSSION

Reliability and Separation

Table 2 shows the reliability and separation indices generated from the Rasch model analysis, which demonstrate the strength and quality of the instrument used to measure pre-service teachers' acceptance and use toward artificial intelligence. The person reliability value of 0.94 indicates that the respondents' responses were highly consistent, suggesting that the instrument can reliably differentiate between individuals with different levels of the measured traits. Similarly, the item reliability of 0.97 reflects that the items themselves are stable and reproducible across different samples, confirming the robustness of the instrument design. The overall Cronbach's alpha of 0.95 further supports excellent internal consistency of the scale. In terms of separation indices, the person separation value of 3.97 suggests that the instrument is capable of categorizing respondents into approximately four distinct ability levels, while the item separation value of 5.65 implies that the sample size was sufficiently large to confirm the hierarchy and difficulty of items, distinguishing around six strata of item difficulty. Finally, the significant chi-square statistic ($\chi^2 = 28{,}447.75$; d.f. = 11,877; $p < .01$) indicates that there are meaningful differences in responses across items, further validating the measurement model. Together, these results confirm that the instrument has excellent reliability, strong discriminatory power, and is statistically sound for use in assessing pre-service teachers' acceptance and use of AI.

Table 2. Reliability and Separation

Indicator	Value
Person Reliability	0.94
Item Reliability	0.97
Cronbach Alpha	0.95
Person Separation	3.97
Item Separation	5.65
Chi-square	28447.75^{**} (d.f. 11877)

Item Fit Statistics

Table 3 presents the item fit statistics of the GAI acceptance and use scale, assessed using the Rasch model through mean square (MNSQ) values for both infit and outfit, as well as point–measure correlation (Pt Mea Corr). Ideally, acceptable MNSQ values range between 0.5 and 1.5, indicating that the items fit well with the underlying measurement model, while point–measure correlations above 0.40 suggest that the items are positively correlated with the overall construct being measured. The results show that all items fall within the recommended MNSQ range, with infit and outfit values generally clustering around 0.60–1.30, demonstrating good model-data fit. For instance, items such as SN1 (Infit = 0.71, Outfit = 0.70) and PE1 (Infit = 0.66, Outfit = 0.66) exhibit excellent fit, while slightly higher outfit values are observed for EE3 (1.24), BI2 (1.19), and AIU2 (1.34), though these still remain within acceptable thresholds. In terms of item discrimination, the Pt Mea Corr values range from 0.58 to 0.80, indicating that all items contribute meaningfully to their respective constructs, with particularly strong correlations observed for AT3 (0.80) and H3 (0.77).

Table 3. Item Fit Statistics of digital skill instrument

Item	MNSQ		Pt Mea Corr
	Infit	Outfit	
SN1	0.71	0.70	0.75
SN2	0.68	0.67	0.74
SN3	0.82	0.82	0.71
SN4	0.70	0.70	0.74
SN5	1.05	1.05	0.62
PE1	0.66	0.66	0.72
PE2	0.75	0.74	0.0.69
PE3	0.79	0.78	0.71
EE1	0.92	0.92	0.62

EE2	1.02	1.01	0.60
EE3	1.24	1.24	0.58
HM1	0.92	0.92	0.67
HM2	0.69	0.62	0.74
HM3	0.70	0.70	0.70
FC1	0.82	0.82	0.65
FC2	0.69	0.62	0.76
FC3	0.61	0.61	0.76
FC4	0.69	0.69	0.73
H1	0.88	0.88	0.70
H2	0.84	0.84	0.68
H3	0.66	0.66	0.77
H4	0.78	0.78	0.75
AT1	0.77	0.77	0.74
AT2	0.61	0.61	0.77
AT3	0.61	0.61	0.80
AT4	0.83	0.82	0.73
PBC1	0.75	0.74	0.76
PBC2	0.76	0.77	0.76
PBC3	1.02	1.01	0.63
BI1	0.87	0.87	0.71
BI2	1.19	1.19	0.70
BI3	0.97	0.97	0.72
AIU1	1.03	1.03	0.71
AIU2	1.28	1.34	0.64
AIU3	1.16	1.20	0.68

Unidimensionality

Table 4 summarizes the unidimensionality test results of the GAI acceptance and use scale using Rasch model analysis. The raw variance explained by persons (17.5%) and items (24.0%) together accounted for a total explained variance of 41.5%, which exceeds the minimum standard of 40% commonly recommended for unidimensional constructs in educational measurement. This indicates that the majority of the variance is captured by the intended latent trait, confirming that the instrument primarily measures a single underlying construct. The unexplained variance in the first contrast yielded an eigenvalue of 1.8, which is below the critical threshold of 2.0, suggesting that no strong secondary dimension exists within the data. Correspondingly, the percentage of unexplained variance in the first contrast was 11.9%, a value that falls within acceptable limits, further supporting the assumption of unidimensionality. Collectively, these findings demonstrate that the digital skill instrument is unidimensional and suitable for measuring pre-service teachers' acceptance and use toward artificial intelligence within a coherent construct.

Table 4. Unidimensionality of digital skill instrument

	Value
Raw variance explained by persons	17.5%
Raw variance explained by items	24.0%
Raw variance explained by measures	41.5%

Unexplained variance in 1st contrast (eigenvalue)	1.8
Unexplained variance in 1st contrast (percentage)	11.9%

The results of the Rasch analysis provide strong evidence that the instrument designed to measure pre-service teachers' acceptance and use of artificial intelligence demonstrates robust psychometric properties and is suitable for educational measurement. High person reliability (0.94) and item reliability (0.97) suggest that both respondents and items behave consistently within the measurement framework, ensuring that the instrument can effectively distinguish between individuals with differing levels of acceptance and usage tendencies. The separation indices further reinforce this finding, indicating that the scale can categorize respondents across multiple strata of ability and confirm a coherent difficulty hierarchy among items. Item fit statistics also show favorable outcomes, with all items falling within acceptable MNSQ ranges and exhibiting strong point–measure correlations, confirming that each item contributes meaningfully to the latent construct. Moreover, the unidimensionality test validates that the instrument primarily measures a single unified construct, as indicated by acceptable explained variance and low residual contrast values. Collectively, these results signify that the instrument is psychometrically sound, with strong reliability, discrimination capacity, and internal coherence, making it suitable for research examining AI-related behavioral tendencies within STEM teacher education.

The psychometric strength demonstrated through the Rasch model has several important implications for research, teacher education, and digital transformation in education. First, the validated instrument offers a rigorous measurement tool for assessing pre-service teachers' readiness to engage with artificial intelligence in instructional contexts, which is critical as AI integration becomes increasingly prominent in STEM learning. Second, the ability of the instrument to differentiate individuals across several strata highlights its applicability for diagnostic purposes, enabling institutions to identify students who may require targeted support or training. Third, the verified unidimensional structure facilitates efficient use of the instrument in large-scale studies and cross-institutional comparisons, supporting the development of evidence-based teacher training curricula. Finally, the instrument may inform policymakers and educational leaders regarding the readiness of future educators to adopt emerging technologies, contributing to strategic planning for AI-enriched learning environments.

CONCLUSION

This study aimed to validate an instrument measuring pre-service STEM teachers' acceptance and use of Artificial Intelligence (AI) through a quantitative, cross-sectional survey involving 401 Generation Z pre-service teachers. Using Rasch modeling, the study examined the psychometric soundness of the instrument across multiple dimensions, including reliability, separation, item fit, and unidimensionality. The results demonstrated that the instrument performs robustly within the Rasch framework, as evidenced by high person and item reliability coefficients, clear separation indices, acceptable mean square fit values, and strong point–measure correlations for all items. The unidimensionality assessment further confirmed that the instrument measures a single coherent construct, indicating structural consistency and internal alignment with its theoretical foundation.

Taken together, the findings affirm that the instrument is both valid and reliable for assessing pre-service teachers' acceptance and use of AI. Beyond its psychometric strength, the instrument contributes to the broader field of STEM teacher education by providing a rigorous measurement tool for understanding AI adoption readiness among future educators. As AI continues to influence teaching and learning practices, such validated instruments are essential for informing teacher preparation programs, guiding institutional interventions, and supporting policy directions related to digital transformation in education. Future research may extend this work by testing the instrument across diverse educational contexts, cultures, and training environments, thereby enhancing its generalizability and further advancing scholarship on AI integration within teacher education.

REFERENCES

Chen, J. J., & Lin, J. C. (2024). Artificial intelligence as a double-edged sword: Wielding the POWER principles to maximize its positive effects and minimize its negative effects. *Contemporary Issues in Early Childhood, 25*(1), 146–153.

Chng, E., Tan, A. L., & Tan, S. C. (2023). Examining the use of emerging technologies in schools: A review of artificial intelligence and immersive technologies in STEM education. *Journal for STEM Education Research, 6*(3), 385–407.

Dewi, H. R., Qudratuddarsi, H., Ningthias, D. P., & Cinthami, R. D. D. (2024). The Current Update of ChatGPT Roles in Science Experiment: A Systemic Literature Review. *Saqbe: Jurnal Sains dan Pembelajarannya*, *1*(2), 74-85.

Guan, L., Zhang, Y., & Gu, M. M. (2025). Pre-service teachers' preparedness for AI-integrated education: An investigation from perceptions, capabilities, and teacher identity changes. *Computers and Education: Artificial Intelligence, 8*, 100341.

Habibi, A., Muhaimin, M., Danibao, B. K., Wibowo, Y. G., Wahyuni, S., & Octavia, A. (2023). ChatGPT in higher education learning: Acceptance and use. *Computers and Education: Artificial Intelligence, 5*, 100190.

Habibi, A., Mukminin, A., Octavia, A., Wahyuni, S., Danibao, B. K., & Wibowo, Y. G. (2024). ChatGPT acceptance and use through UTAUT and TPB: A big survey in five Indonesian universities. *Social Sciences & Humanities Open, 10*, 101136.

Hidayat, R., Qudratuddarsi, H., Ayub, A. F. M., & Latif, I. N. A. (2025). Psychometric Properties Of The Perma-Profiler For Indonesian College Students: A Rasch Modelling Analysis. *Journal of Institutional Research South East Asia*, *23*(1).

Hope, D., Kluth, D., Homer, M., Dewar, A., Goddard-Fuller, R., Jaap, A., & Cameron, H. (2025). Exploring the use of Rasch modelling in "common content" items for multi-site and multi-year assessment. *Advances in Health Sciences Education*, *30*(2), 427-438.

Jamieson, M. K., Govaart, G. H., & Pownall, M. (2023). Reflexivity in quantitative research: A rationale and beginner's guide. *Social and Personality Psychology Compass*, *17*(4), e12735.

Kesmodel, U. S. (2018). Cross-sectional studies–what are they good for?. *Acta obstetricia et gynecologica Scandinavica*, *97*(4), 388-393.

Laru, J., Celik, I., Jokela, I., & Mäkitalo, K. (2025). The antecedents of pre-service teachers' AI literacy. *European Journal of Teacher Education*, 1–23.

Medvedev, O. N., & Krägeloh, C. U. (2025). Rasch measurement model. In *Handbook of assessment in mindfulness research* (pp. 131-147). Cham: Springer Nature Switzerland.

Ouyang, F., Dinh, T. A., & Xu, W. (2023). A systematic review of AI-driven educational assessment in STEM education. *Journal for STEM Education Research, 6*(3), 408–426.

Özden, M., Yaşar, F. Ö., & Meydan, E. (2025). The relationship between pre-service teachers' attitude towards artificial intelligence and their AI literacy. *Pegem Journal of Education and Instruction, 15*(3), 121–131.

Parviz, M. (2024). AI in education: Comparative perspectives from STEM and non-STEM instructors. *Computers and Education Open, 6*, 100190.

Plonsky, L. (2017). Quantitative research methods. In *The Routledge handbook of instructed second language acquisition* (pp. 505-521). Routledge.

Qudratuddarsi, H., Fauziah, A., Agung, A., & Yanti, M. (2025). "Status Quo" ChatGPT dalam Pengajaran dan Pembelajaran Fisika: Systematic Literature Review. *PHY*, 110-118.

Saddia, A., Yanti, M., & Qudratuddarsi, H. (2025). The impact of chatgpt-assisted problem-based learning on students physics achievement and their chatgpt acceptance. *Compton: Jurnal Ilmiah Pendidikan Fisika*, *12*(1), 42-56.

Schiff, D. (2022). Education for AI, not AI for education. *International Journal of Artificial Intelligence in Education, 32*(3), 527–563.

Sun, F., Tian, P., Sun, D., Fan, Y., & Yang, Y. (2024). Pre-service teachers' inclination to integrate AI into STEM education. *British Journal of Educational Technology, 55*(6), 2574–2596.

Wang, D., & Huang, X. (2025). Transforming education through artificial intelligence. *Interactive Learning Environments*, 1–20.

Xu, W., & Ouyang, F. (2022). Application of AI technologies in STEM education. *International Journal of STEM Education, 9*(1), 59.

Yang, Y., Sun, W., Sun, D., & Salas-Pilco, S. Z. (2025). Navigating the AI-enhanced STEM education landscape. *Research in Science & Technological Education, 43*(3), 693–717.

Yang, Y., Zhang, Y., Sun, D., He, W., & Wei, Y. (2025). Navigating the landscape of AI literacy education. *Humanities and Social Sciences Communications, 12*(1), 1–12.

Zhai, X., Neumann, K., & Krajcik, J. (2023). AI for tackling STEM education challenges. *Frontiers in Education, 8*, 1183030.

Zhang, C., Schießl, J., Plößl, L., Hofmann, F., & Gläser-Zikuda, M. (2023). Acceptance of artificial intelligence among pre-service teachers. *International Journal of Educational Technology in Higher Education, 20*(1), 49.