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

Saqbe : Sains dan Pembelajarannya

Validation of Artificial Intelligence Expectations and Training Attitudes (AI-ETA) Scale for Pre-service Chemistry Teachers

Validasi Instrumen untuk Mengukur Artificial Intelligence Expectations and Training Attitudes (AI-ETA) pada Calon Guru Kimia

Dea Santika Rahayu[1], Syamsul Alim Bahri[2], Rustam Effendy Simamora[3]

[1]Pendidikan IPA, Fakultas Keguruan dan Ilmu Pendidikan, Universitas Tidar
[2]Hukum keluarga, Institut Agama Islam As'adiyah Sengkang
[3]Pendidikan Matematika, Fakultas Keguruan dan Ilmu Pendidikan, Universitas Borneo Tarakan
*Correspondence e-mail: dea.santika@untidar.ac.id

Received: 15 July 2025
Revised: 22 August 2025
Accepted : 10 September 2025
Published: 30 September 2025

Abstrak

Artificial intelligence (AI) is increasingly integrated into education, requiring valid instruments to assess teachers' expectations and training attitudes toward its use. This study aims to validate the Artificial Intelligence Expectations and Training Attitudes (AI-ETA) scale for pre-service chemistry teachers. A quantitative survey design was employed, involving 95 participants from two public universities. The instrument consisted of two dimensions: expectations about AI's impact on education and expectations about AI training courses, measured using a five-point rating scale. Data were analyzed using the Rasch measurement model to examine item fit, reliability, separation, and unidimensionality. The results indicate that the instrument demonstrates strong psychometric properties, with high person reliability and adequate item separation. Most items showed acceptable fit statistics, although a few items required minor refinement. The unidimensionality analysis confirmed that the scale measures a single dominant construct. These findings suggest that the AI-ETA scale is a valid and reliable instrument for assessing pre-service teachers' perceptions of AI. The study highlights the importance of understanding teachers' expectations to support effective AI integration in chemistry education.

Keyword

Artificial Intelligence (AI) in education, Instrument validation, Pre-service teachers' attitudes and expectations, Rasch model, Chemistry education

Abstract

Kecerdasan buatan (AI) semakin banyak diintegrasikan dalam pendidikan, sehingga diperlukan instrumen yang valid untuk mengukur ekspektasi dan sikap calon guru terhadap pelatihan AI. Penelitian ini bertujuan untuk memvalidasi skala Artificial Intelligence Expectations and Training Attitudes (AI-ETA) pada calon guru kimia. Penelitian menggunakan pendekatan kuantitatif dengan desain survei yang melibatkan 95 responden dari dua universitas negeri. Instrumen terdiri dari dua dimensi, yaitu ekspektasi terhadap dampak AI dalam pendidikan dan ekspektasi terhadap pelatihan AI, yang diukur menggunakan skala lima poin. Analisis data dilakukan menggunakan model Rasch untuk menguji kesesuaian item, reliabilitas, separasi, dan unidimensionalitas. Hasil penelitian menunjukkan bahwa instrumen memiliki kualitas psikometrik yang baik, ditandai dengan reliabilitas tinggi dan separasi item yang memadai. Sebagian besar item memenuhi kriteria fit, meskipun beberapa item memerlukan perbaikan. Analisis unidimensionalitas menunjukkan bahwa instrumen mengukur satu konstruk utama. Temuan ini menunjukkan bahwa skala AI-ETA valid dan reliabel untuk digunakan dalam konteks pendidikan kimia serta mendukung integrasi AI secara efektif.

Kata Kunci

Artificial Intelligence (AI) dalam Pendidikan, Validasi instrument, Sikap dan ekspektasi guru prajabatan, Model Rasch, Pendidikan kimia

INTRODUCTION

Artificial intelligence (AI) is increasingly used and expected to play a transformative role in teaching and learning, including in chemistry education. AI enables adaptive learning environments that tailor instruction to individual student needs, allowing learners to progress at their own pace while addressing misconceptions in real time. In chemistry, AI-powered simulations and virtual laboratories provide safe, cost-effective alternatives to traditional experiments, helping students visualize abstract concepts such as molecular interactions and reaction mechanisms (Erümit & Sarıalioğlu, 2025; Iyamuremye et al., 2024). Intelligent tutoring systems can guide students through problem-solving processes, offering immediate feedback and scaffolding support. Additionally, AI can assist teachers by automating administrative tasks such as grading and assessment, freeing time for more meaningful instructional interactions. Learning analytics derived from AI systems can also provide insights into student performance and engagement, enabling data-driven instructional decisions. As digital transformation accelerates, AI is expected to become an integral component of modern education, reshaping both pedagogy and classroom practices, particularly in complex and concept-heavy subjects like chemistry (Lizano-Sánchez, et al., 2025; Naeyaert, Scarratt, Murphy & Pullen, 2025).

AI can be used in many ways to enhance chemistry teaching. It supports virtual labs, allowing students to conduct experiments safely and repeatedly. AI-powered tutors can guide learners through chemical calculations and conceptual questions step by step. Automated assessment tools can quickly evaluate assignments and provide feedback. AI also enables visualization tools that help students understand molecular structures and chemical reactions. Furthermore, predictive analytics can identify students at risk of misunderstanding key topics. By integrating these tools, teachers can make chemistry more interactive, personalized, and accessible, improving both engagement and learning outcomes (Kaana, Ugosor, & Ukpoko, 2026; Wright, Diaz, & von Gillern, 2026).

The integration of AI in education largely depends on teachers, including pre-service teachers, as they are the primary agents implementing technology in classrooms. Their beliefs, knowledge, and confidence in using AI significantly influence how effectively these tools are adopted. Pre-service teachers, in particular, are in a critical phase of developing their pedagogical and technological competencies. If they are equipped with positive attitudes and sufficient understanding of AI, they are more likely to integrate it meaningfully into their future teaching practices. However, their readiness is shaped by their training experiences, exposure to digital tools, and opportunities to practice using AI in instructional contexts. Without proper preparation, teachers may resist or underutilize AI technologies. Therefore, empowering both in-service and pre-service teachers with the necessary skills and mindset is essential for successful AI integration in education, especially in specialized subjects like chemistry (Kashif, Ammar, Sellami, Chiu, Abbasi, & Ahmad, 2025; Watted, 2025).

Despite the growing importance of AI in education, training opportunities for teachers, including pre-service teachers, remain limited in many educational programs. Teacher education curricula often focus more on general pedagogical skills and traditional educational technologies, with insufficient emphasis on emerging tools such as AI. As a result, many teachers graduate without a clear understanding of how AI can be applied in teaching and learning contexts. Even when training is available, it may be fragmented, overly theoretical, or lacking in practical application. Professional development programs also tend to offer limited exposure to AI, making it difficult for teachers to stay updated with rapidly evolving technologies. This gap can lead to uncertainty, lack of confidence, and resistance toward adopting AI in classrooms. Furthermore, without structured training that includes ethical considerations, practical implementation, and real-world examples, teachers may struggle to use AI effectively. Expanding and improving AI-focused training in teacher education is therefore essential to prepare educators for future demands (Guan, Zhang, & Gu, 2025; Xue, Ghazali, & Mahat, 2025).

Understanding teachers' expectations and attitudes toward AI is essential because these factors directly influence whether and how AI is adopted in educational practice. Teachers' beliefs about the potential of AI to transform teaching, enhance student learning, and support instructional tasks shape their willingness to engage with these technologies (Khan, Kausar, & Khan, 2025; Marsalek, & Teplá, 2026). Positive expectations can encourage experimentation, innovation, and integration into classroom activities, while uncertainty or negative perceptions may lead to resistance or superficial use. Exploring how teachers perceive AI's impact on student outcomes, engagement, and classroom dynamics provides insight into their readiness to embrace change. This understanding is particularly important in anticipating how educators envision the role of AI in the near future and how it may align with their pedagogical goals and teaching contexts (Yehya, ElSayary, Al Murshidi, & Al Zaabi, 2025).

Equally important is understanding teachers' expectations regarding their preparation to use AI effectively. Their experiences with training, as well as their perceived needs for knowledge, skills, and support, influence their confidence and competence in applying AI in educational settings (Tomczyk, & Majkut, 2025). Teachers may differ in the extent to which they value practical experience, conceptual understanding, ethical awareness, and opportunities to apply learning in authentic contexts. Identifying these expectations helps reveal gaps between existing professional development provisions and what teachers actually require. Furthermore, understanding the types of AI tools teachers are familiar with or interested in learning can inform the design of more relevant and responsive training programs (Daher, 2025; Khalid, Ifran, Huang, Tadesse, & Dainkun, 2026). Although numerous studies highlight the potential of AI in education, there is a limited understanding of how pre-service teachers perceive AI and what they expect from its integration into teaching and learning, particularly in subject-specific contexts like chemistry. Existing research often focuses on technological capabilities or general teacher readiness, rather than examining detailed expectations about AI applications and training experiences. Furthermore, there is insufficient attention to the mismatch between available AI training and teachers' actual needs, including practical implementation, ethical awareness, and real-world applications. This gap is significant because pre-service teachers represent the future workforce responsible for adopting AI in classrooms. Investigating their attitudes and expectations can provide deeper insights into barriers and opportunities, ultimately informing the design of more effective teacher education programs and AI integration strategies.

METHOD

Research Design

This study adopted a quantitative survey design because the main purpose was to validate Artificial Intelligence Expectations and Training Attitudes (AI-ETA) for pre-service chemistry teachers using numerical evidence. A survey method was considered appropriate since validation requires collecting responses from a relatively large group in a standardized manner to evaluate item performance objectively. The two-tier format was selected to capture not only participants' answer choices but also the reasoning behind those choices, allowing the identification of misconceptions and levels of scientific justification. Integrating ethnochemistry into the items required empirical testing to ensure that culturally contextualized content remained measurable, unbiased, and interpretable across respondents (Fischer, Boone, & Neumann, 2023). Quantitative analysis of survey data using the Rasch measurement model enabled systematic examination of item functioning, including item difficulty calibration, person ability estimation, and item fit statistics, providing defensible evidence of construct validity and reliability. The Rasch model also supported evaluation of unidimensionality and separation indices, ensuring that the ethnochemistry-integrated two-tier instrument operates consistently and measures chemical literacy accurately among pre-service chemistry teachers (Qudratuddarsi, Meivawati, & Saputra, 2024; Schoonenboom, 2023).

Research Subject

Convenience sampling was used because the research focused on initial instrument validation with accessible respondents who matched the target characteristics and were available during the data collection period. This technique allowed efficient recruitment within limited time and resources while ensuring participants had sufficient background in chemistry education to respond meaningfully to the two-tier items (Golzar, Noor, & Tajik, 2022; Qudratuddarsi, Rahmah, & Indriyanti, 2025). The participants consisted of 95 pre-service chemistry teachers enrolled in two public universities in West Nusa Tenggara, selected due to their relevance as prospective chemistry educators and their exposure to chemistry concepts and local cultural contexts reflected in the instrument.

Instrument

The instrument used in this study was adapted from Reina-Parrado, Román-Graván, and Hervás-Gómez (2025) and was subjected to a back-translation process to ensure linguistic accuracy and conceptual equivalence between the original and translated versions. Back-translation was applied to minimize misinterpretation and maintain the validity of the instrument across different language contexts (Edunov, Ott, Auli & Grangier, 2018). The instrument consists of two main dimensions: Expectations about AI and Its Impact on Education (EAII) and Expectations about AI Training Courses (EAITC). Each item is measured using a five-point rating scale, allowing for the capture of variations in respondents' perceptions while ensuring clarity, consistency, and ease of analysis. This structure enables a comprehensive assessment of pre-service teachers' expectations and attitudes toward AI in educational contexts.

Data Collection and Analysis

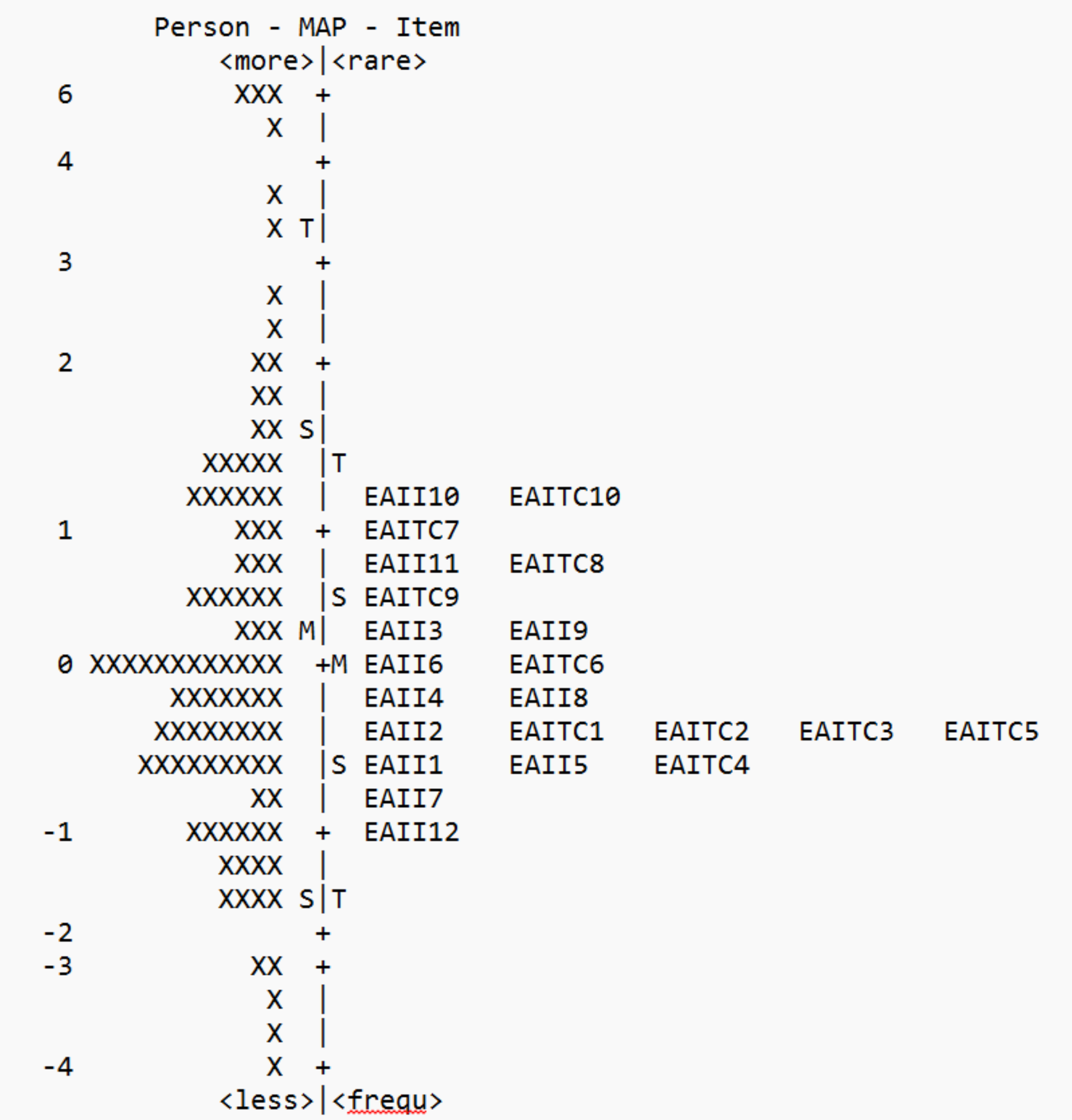
Data collection was conducted through Google Forms, supporting sustainable, paperless research practices while enhancing efficiency and minimizing data entry errors. The digital format also enabled real-time access to responses (Saddia, Yanti, & Qudratuddarsi, 2025). To ensure clarity and accurate understanding of the survey items, the researcher remained present during the data collection process, providing immediate assistance when needed. This approach helped create a supportive environment that encouraged participants to respond sincerely. Participation was voluntary, and respondents were assured that their answers would remain confidential and would not affect their academic evaluation (Jumriani, et al., 2025). These ethical considerations were essential for maintaining the credibility and integrity of the collected data.

Data analysis employed the Rasch measurement model to examine the psychometric quality of the ethnochemistry-integrated two-tier instrument. Responses were coded and entered into Rasch analysis software to estimate item difficulty and person ability on a common logit scale. The analysis included examination of item-person distribution through a Wright map, reliability and separation indices for persons and items, and internal consistency using Cronbach's alpha. Item fit was evaluated using Infit and Outfit mean square (MNSQ) statistics, standardized fit indices (ZSTD), and point-measure correlations to confirm alignment with the intended construct. Unidimensionality was assessed using principal component analysis (PCA) of standardized residuals (Adam, Qudratuddarsi, Ningthias, Rahmadhani, & Noviana, 2025; Hidayat, Qudratuddarsi, Ayub & Latif, 2025).

RESULT AND DISCUSSION

Wright Map

The Wright map is a visual representation produced through Rasch model analysis that places respondents' abilities and item difficulties on the same logit scale, allowing direct comparison between learner competence and assessment demands (Ismail, Din, & Jusoh, 2021; Sutrisno, Manuharawati, & Masriyah, 2025). The Wright map (Person–Item Map) in Fugure 1 shows that the distribution of respondents' abilities is generally centered around the mean (0 logits), with most pre-service chemistry teachers clustered between 0 and +2 logits, indicating moderate to high expectations toward AI. The item distribution appears well-targeted, as most items are located within this range, suggesting good alignment between item difficulty and respondent ability. Items such as EAII10, EAITC10, and EAITC8 are positioned at higher logit levels, indicating they are more difficult to endorse and reflect higher-order expectations about AI impact and training outcomes. In contrast, items like EAII1, EAII5, EAITC1, and EAITC2 are located at lower logit levels, suggesting they are easier to agree with and represent more general or commonly accepted perceptions. The spread of items across logits indicates adequate coverage of the construct, although a slight concentration around the middle suggests moderate differentiation. Overall, the instrument demonstrates acceptable targeting and construct representation for measuring AI expectations and training attitudes.



```
       Person - MAP - Item
           <more>|<rare>
 6          XXX  +
              X  |
 4               +
              X  |
              X T|
 3               +
              X  |
              X  |
 2           XX  +
             XX  |
             XX S|
          XXXXX  |T
         XXXXXX  |  EAII10   EAITC10
 1          XXX  +  EAITC7
            XXX  |  EAII11   EAITC8
         XXXXXX  |S EAITC9
            XXX M|  EAII3    EAII9
 0 XXXXXXXXXXXX  +M EAII6    EAITC6
        XXXXXXX  |  EAII4    EAII8
       XXXXXXXX  |  EAII2    EAITC1   EAITC2   EAITC3   EAITC5
      XXXXXXXXX  |S EAII1    EAII5    EAITC4
             XX  |  EAII7
-1       XXXXXX  +  EAII12
           XXXX  |
           XXXX S|T
-2               +
-3           XX  +
              X  |
              X  |
-4            X  +
           <less>|<frequ>
```

Figure 1. Wright Map

Reliability and Separation

In Rasch model analysis, reliability and separation indices are used to evaluate the consistency of measurement and the instrument's ability to distinguish between different levels of respondent ability and item difficulty. Person reliability reflects how consistently the instrument differentiates respondents based on chemical literacy, while item reliability indicates the stability of item difficulty estimates across samples

(Medvedev & Krägeloh, 2025). The reliability and separation indices indicate that the AI-ETA scale demonstrates strong psychometric properties in measuring pre-service chemistry teachers' expectations and training attitudes toward artificial intelligence. The person reliability value of 0.93 suggests high internal consistency among respondents, indicating that the instrument can reliably distinguish between individuals with different levels of AI-related expectations. The person separation index of 0.94 further confirms that the sample can be divided into distinct strata based on their responses. On the other hand, the item separation value of 3.78 and item reliability, which is expected to be high, indicate that the items are well distributed along the latent trait and can effectively differentiate varying levels of construct difficulty. The Cronbach's alpha value, although partially reported, supports the overall consistency of the instrument. Additionally, the significant chi-square value (4121.272, $p < 0.01$) indicates that the model fits the data well and that item difficulties vary meaningfully.

Table 1. Reliability and Separation

	Value
Person Reliability	0.93
Person Separation	0.94
Item Reliability	3.93
Item Separation	3.78
Cronbach alpha	0.
Chi-square	4121.272** with 296

Item Fit Statistics

Item fit statistics in the Rasch model are used to determine whether each item functions as expected in measuring a single latent trait, which in this study is chemical literacy among preservice chemistry teachers through ethnochemistry-integrated two-tier multiple-choice questions. Fit is commonly evaluated using Infit and Outfit mean square (MNSQ) values and their standardized forms (ZSTD). Infit is more sensitive to unexpected responses close to a respondent's ability level, whereas Outfit is influenced by outliers and unexpected responses far from the ability estimate. Generally, MNSQ values between 0.5–1.5 and ZSTD values within ±2 indicate acceptable fit. Positive point-measure correlations (Pt Mea Corr) further confirm that items align with the intended construct.

The item fit statistics indicate that most items in the AI-ETA scale fall within the acceptable range for Rasch model fit, typically between 0.5 and 1.5 for mean square (MNSQ) values. Items such as EAII3, EAII4, EAII8, EAII9, EAII12, and several EAITC items demonstrate good fit, with infit and outfit values close to 1.0, suggesting that these items function consistently with model expectations. However, some items, including EAII5, EAII10, EAII11, and EAITC8, show slightly elevated MNSQ values above 1.3, indicating potential noise or unpredictability in responses. In particular, EAII11 presents the highest misfit, with infit and outfit values exceeding 1.7, suggesting that this item may not align well with the underlying construct. Conversely, items such as EAII1, EAII2, EAITC4, and EAITC5 display lower MNSQ values, indicating possible redundancy or overpredictability in responses.

The standardized fit statistics (ZSTD) further support these findings, where most items fall within the acceptable range of -2 to +2, although some items such as EAII11, EAII2, and EAITC8 exceed these thresholds, reinforcing concerns about item misfit. The point-measure correlations (Pt Mea Corr) for all items are positive and range from 0.52 to 0.81, indicating that each item contributes meaningfully to the overall construct measurement. Higher correlations, such as those found in EAII2 and EAITC5, suggest strong alignment with the latent trait, while relatively lower values, such as EAII11 and EAITC8, indicate weaker associations. Overall, the pattern suggests that the majority of items are valid indicators of AI expectations and training attitudes, with a few items requiring further review or refinement.

Table 2. Item Fit Statistics

No	Item	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD	Pt Mea Corr
1	EAII1	0.6	-3.0	0.61	-2.6	0.78
2	EAII2	0.51	-3.9	0.5	-3.5	0.81
3	EAII3	1.11	0.7	1.02	0.2	0.68
4	EAII4	0.84	-1.0	0.83	-1.0	0.74
5	EAII5	1.35	2.1	1.41	2.2	0.65
6	EAII6	0.76	-1.6	0.76	-1.5	0.78
7	EAII7	0.78	-1.6	0.75	-1.5	0.73
8	EAII8	0.95	-0.3	0.95	-0.2	0.75
9	EAII9	1.1	0.7	1.06	0.4	0.66
10	EAII10	1.37	2.4	1.39	2.3	0.67
11	EAII11	1.7	4.0	1.75	3.9	0.52
12	EAII12	1.11	0.8	1.1	0.6	0.66
13	EAITC1	0.81	-1.2	0.78	-1.3	0.72
14	EAITC2	0.81	-1.2	0.79	-1.2	0.74
15	EAITC3	1.24	1.5	1.19	1.1	0.66
16	EAITC4	0.7	-2.2	0.68	-2.0	0.76
17	EAITC5	0.61	-2.9	0.58	-2.8	0.81
18	EAITC6	0.73	-1.9	0.65	-2.3	0.78
19	EAITC7	1.14	1.0	1.13	0.8	0.74
20	EAITC8	1.39	2.4	1.52	2.9	0.63
21	EAITC9	1.05	0.4	1.09	0.6	0.70
22	EAITC10	1.0	0.1	1.03	0.3	0.70

Unidimensionality

Unidimensionality in the Rasch model refers to the assumption that an instrument measures a single dominant latent trait, ensuring that item responses are primarily driven by one construct. This assumption is commonly evaluated using principal component analysis (PCA) of standardized residuals, where the proportion of variance explained by Rasch measures indicates the strength of the intended dimension. The standardized residual variance analysis indicates that the AI-ETA scale demonstrates adequate unidimensionality and construct representation. The total raw variance in observations is 47.7 (100%), reflecting the overall variability captured in the dataset. Of this, 25.7 (53.9%) of the variance is explained by the Rasch measures, which closely aligns with the modeled expectation of 53.4%, suggesting a strong correspondence between empirical data and the measurement model. This proportion exceeds the commonly accepted threshold of 40%, indicating that the instrument sufficiently captures the primary latent construct of artificial intelligence expectations and training attitudes. Furthermore, the variance explained by persons is 14.9 (31.3%), showing that individual differences among respondents contribute substantially to the measured construct. Meanwhile, the variance explained by items is 10.8 (22.6%), indicating that item difficulty is well distributed across the scale. The similarity between empirical and modeled values supports the structural validity of the instrument.

Table 3. Standardized Residula Variance

	Empirical (Eigenvalue)	Empirical (%)	Modeled (%)
Total raw variance in observations	47.7	100	100
Raw variance explained by measures	25.7	53.9	53.4
Raw variance explained by persons	14.9	31.3	31.0
Raw variance explained by items	10.8	22.6	22.4

The findings of this study indicate that the AI-ETA scale demonstrates strong psychometric quality in assessing pre-service chemistry teachers' expectations and training attitudes toward artificial intelligence. The Rasch analysis shows that the instrument is well-targeted, with item difficulties aligned to respondents' ability levels, and most items functioning within acceptable fit criteria. Reliability and separation indices confirm that the scale can consistently distinguish between different levels of expectations among respondents. Although a small number of items exhibit minor misfit, the overall structure remains stable and coherent. The unidimensionality results further support that the instrument measures a single dominant construct, ensuring construct validity and interpretability of the results.

Despite these strengths, several limitations should be considered. The sample size is relatively small and limited to two universities, which may restrict the generalizability of the findings to broader populations. The use of convenience sampling may also introduce bias, as participants may not fully represent the diversity of pre-service teachers. Additionally, the study relies solely on self-reported data, which may be influenced by social desirability or response bias. Some items that showed misfit suggest that further refinement is needed to improve measurement precision. The cross-sectional design also limits the ability to examine changes in expectations and attitudes over time.

The implications of this study are significant for both research and practice in AI integration in education. The validated AI-ETA scale provides a reliable tool for measuring pre-service teachers' expectations and training needs, which can inform the design of more effective teacher education programs. Educational institutions can use the findings to develop targeted AI training that aligns with teachers' perceived needs, including practical applications, ethical considerations, and real-world examples. For researchers, the instrument offers a foundation for further studies examining the relationship between expectations, attitudes, and actual AI adoption in teaching. Additionally, policymakers can utilize these insights to support the development of strategies that promote meaningful and sustainable integration of AI in chemistry education.

CONCLUSION

The results of this study indicate that the Artificial Intelligence Expectations and Training Attitudes (AI-ETA) scale demonstrates strong psychometric properties in measuring pre-service chemistry teachers' perceptions of AI in education. The Wright map shows a well-targeted distribution between person ability and item difficulty, indicating that the instrument appropriately captures varying levels of expectations. Reliability and separation indices further confirm the consistency and discriminative power of the scale, with high person reliability and adequate item separation suggesting that the instrument can distinguish between different respondent levels effectively. Item fit statistics reveal that most items function well within the Rasch model expectations, although a few items exhibit slight misfit and may require refinement. Additionally, the unidimensionality analysis supports that the instrument measures a single dominant construct, as the variance explained by the model exceeds the recommended threshold, indicating structural validity.

Future research should focus on refining items that demonstrate misfit, particularly those with higher mean square values or lower point-measure correlations, to improve overall instrument precision. Expanding the sample size and including participants from diverse educational contexts would enhance the generalizability of the findings. Further studies could also examine the predictive validity of the AI-ETA scale by linking expectations and attitudes toward actual AI integration practices in teaching. Additionally, integrating qualitative approaches such as interviews or open-ended responses may provide deeper insights into the reasons behind participants' responses and help improve item wording. Longitudinal studies are also recommended to explore how pre-service teachers' expectations and training attitudes evolve over time with increased exposure to AI technologies and training programs in educational settings.

REFERENCES

Adam, W., Qudratuddarsi, H., Ningthias, D. P., Rahmadhani, A., & Noviana, E. (2025). Validation of pre-service science teacher artificial intelligence competence self-efficacy (AICS): Rasch model analysis. *Jurnal Ilmiah Profesi Pendidikan*, *10*(2), 1985-1995.

Daher, R. (2025). Integrating AI literacy into teacher education: a critical perspective paper. *Discover Artificial Intelligence*, *5*(1), 217.

Edunov, S., Ott, M., Auli, M., & Grangier, D. (2018). Understanding back-translation at scale. In *Proceedings of the 2018 conference on empirical methods in natural language processing* (pp. 489-500).

Erümit, A. K., & Sarıalioğlu, R. Ö. (2025). Artificial intelligence in science and chemistry education: a systematic review. *Discover Education*, *4*(1), 178.

Fischer, H. E., Boone, W. J., & Neumann, K. (2023). Quantitative research designs and approaches. In *Handbook of research on science education* (pp. 28-59). Routledge.

Golzar, J., Noor, S., & Tajik, O. (2022). Convenience sampling. *International Journal of Education & Language Studies*, *1*(2), 72-77.

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

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).

Ismail, M. S., Din, M. H., & Jusoh, M. S. (2021, July). Predictive modelling using Rasch's person-item map: Intrepreting and assessing in Malaysia manufacturing firms. In *AIP Conference Proceedings* (Vol. 2347, No. 1, p. 020288). AIP Publishing LLC

Iyamuremye, A., Niyonzima, F. N., Mukiza, J., Twagilimana, I., Nyirahabimana, P., Nsengimana, T., ... & Nsabayezu, E. (2024). Utilization of artificial intelligence and machine learning in chemistry education: a critical review. *Discover Education*, *3*(1), 95.

Jumriani, J., et al. (2025). Validation of instrument to measure Gen-Z pre-service teacher chemistry virtual lab acceptance and use. *Chemistry Education Practice,* *8*(1), 128–136. https://doi.org/10.29303/cep.v8i1.8701

Kaana, A., Ugosor, T. P., & Ukpoko, W. U. (2026). Chemistry Education and Artificial Intelligence (AI) In the 21st Century Nigeria: Prospects, Challenges, and the Way Forward. A Review Article. *International Journal of Research and Innovation in Social Science (IJRISS)*, *9*(12).

Khalid, S., Ifran, S., Huang, X., Tadesse, E., & Dainkun, J. (2026). AI-Integrated Professional Development for Pre-Service Teachers: A Systematic Review of Impacts on Pedagogical Strategies and Technological Self-Efficacy. *European Journal of Education*, *61*(2), e70547.

Khan, H. A., Kausar, T., & Khan, S. (2025). Leading AI-Integrated Pedagogies: A Qualitative Study on How School Leaders Guide Teachers in Using AI Tools for Teaching and Learning in Pakistan. *Social Science Review Archives*, *3*(4), 3086-3101.

Kashif, M., Ammar, M., Sellami, A., Chiu, T. K., Abbasi, S. A., & Ahmad, Z. (2025). Teachers' perspectives on AI integration in K-12 education: challenges, opportunities, and preliminary assessment model–a systematic review. *Computers in the Schools*, 1-27.

Lizano-Sánchez, F., Idoyaga, I. J., Orduna, P., Rodriguez-Gil, L., & Arguedas-Matarrita, C. (2025, November). Students' interactions with an artificial intelligence assistant in a remote chemistry laboratory. In *Frontiers in Education* (Vol. 10, p. 1712743). Frontiers Media SA.

Marsalek, R., & Teplá, M. (2026). "Is AI Inevitable?" Development of Attitudes and Practices of Czech Teachers Between 2023 and 2025. *Education Sciences*, *16*(2), 335.

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

Naeyaert, P. J., Scarratt, L. R., Murphy, T., & Pullen, R. (2025). Think–Pair–Chatbot–Share: AI-Facilitated Peer Learning in Chemistry. *Journal of Chemical Education*, *102*(9), 3936-3944.

Qudratuddarsi, H., Meivawati, E., & Saputra, R. (2024). Pelatihan Penelitian Metode Kuantitatif dan Systematic Literature Review bagi Dosen dan Mahasiswa. *Beru'-beru': Jurnal Pengabdian kepada Masyarakat*, *3*(1), 22-32.

Qudratuddarsi, H., Rahmah, N., & Indriyanti, N. (2025). Analysis of the attitudes of generation z prospective science teachers toward the chemistry virtual laboratory. *Arfak Chem: Chemistry Education Journal*, *8*(1), 651-665.

Reina-Parrado, M., Román-Graván, P., & Hervás-Gómez, C. (2025). Learning to teach AI: design and validation of a questionnaire on artificial intelligence training for teachers. *European Journal of Educational Research*, *14*(1), 249-265.

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. https://doi.org/10.30738/cjipf.v12i1.19852

Schoonenboom, J. (2023). The fundamental difference between qualitative and quantitative data in mixed methods research. In *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research* (Vol. 24, No. 1). DEU.

Sutrisno, Manuharawati, & Masriyah. (2025, November). Assessing beliefs about statistics of prospective mathematics teachers using the rasch wright map. In *AIP Conference Proceedings* (Vol. 3333, No. 1, p. 020007). AIP Publishing LLC.

Tomczyk, Ł., & Majkut, A. (2025). Integrating AI in education: An analysis of factors influencing the acceptance, concerns, attitudes, competencies and use of generative artificial intelligence among Polish teachers. *Human Behavior and Emerging Technologies*, *2025*(1), 5599169.

Watted, A. (2025). Teachers' perceptions and intentions toward AI integration in education: insights from the UTAUT model. *Power System Technology*, *49*(2), 164-183.

Wright, G. W., Diaz, B., & von Gillern, S. (2026). Science Educators' Attitudes and Perspectives on Artificial Intelligence (SEAP-AI) Scale. *Journal of Science Education and Technology*, 1-14.

Xue, L., Ghazali, N., & Mahat, J. (2025). Artificial Intelligence (AI) Adoption among Teachers: A Systematic Review and Agenda for Future Research. *International Journal of Technology in Education*, *8*(3), 802-824.

Yehya, F., ElSayary, A., Al Murshidi, G., & Al Zaabi, A. (2025). Artificial intelligence integration and teachers' self-efficacy in physics classrooms. *Eurasia Journal of Mathematics, Science and Technology Education*, *21*(8), em2679.