PANYONARA: Journal of English Education
Vol. 7, No. 2, September 2025
DOI: 10.19105/panyonara.v7i2.21416

ARE PRE-SERVICE EFL TEACHERS READY FOR AI-ASSISTED ASSESSMENT? THE ROLE OF ASSESSMENT LITERACY IN THE DIGITAL ERA

Vina Ayuningtryas[1]
[1]Faculty of Cultural Studies, Brawijaya University, Malang, Indonesia
(vinayuningtryas@student.ub.ac.id)

Ive Emaliana[2]
[2]Faculty of Cultural Studies, Brawijaya University, Malang, Indonesia
(ive@ub.ac.id)

Abstract: *Artificial Intelligence (AI) is transforming educational assessment, particularly in English as a Foreign Language (EFL) settings. As the use of AI is becoming increasingly rapid, pre-service teachers must become proficient with both Language Assessment Literacy (LAL) and preparation to utilize AI tools. Therefore, this study explored the relationship between LAL and AI-based assessment readiness among Indonesian EFL pre-service teachers. Adopting a mixed-method explanatory sequential research design, 200 respondents across 60 universities in Indonesia participated through a questionnaire survey followed by interviews with the highest and lowest AI readiness scores. The quantitative data found a significant moderate positive correlation, implying that higher LAL is associated with greater readiness to use AI-based assessment. Moreover, the descriptive data indicated that while most participants demonstrated high LAL, their AI readiness was only moderately high. Qualitative data revealed that the respondents with better LAL have critical views about practices when it comes to assessment. The research concluded that to effectively integrate AI into assessment practices, pre-service teachers not only need technology skills training but also a solid assessment knowledge. These results have implications for the curriculum in teacher education, for which there is demand for integrated frameworks that link assessment theory with ethical AI implementation.*

Keywords: *Language Assessment Literacy; AI-Based Assessment Readiness; EFL*

Article History:

Received	: 05 August 2025	Accepted	: 30 September 2025
Revised	: 14 September 2025	Available Online	: 30 September 2025

Suggested Citation Format:
Ayuningtryas, V., & Emaliana, I. (2025). Are Pre-Service EFL Teachers Ready for AI-Assisted Assessment? The Role of Assessment Literacy in the Digital Era. *PANYONARA: Journal of English Education, 7*(2), 267–294. https://doi.org/10.19105/panyonara.v7i2.21416

INTRODUCTION

Artificial Intelligence (AI) has emerged as a transformative force in education, revolutionizing how teachers instruct and assess learners. The integration of AI in English as a Foreign Language (EFL) education, particularly in language assessment, is gaining increasing attention due to its potential to enhance both teaching efficacy and student

engagement. Rapid technological advancement has facilitated the development of AI tools that support automated grading, adaptive feedback, and personalized learning pathways. Digital technology, including AI, presents vast opportunities to improve pedagogical practices in EFL classrooms (Mekheimer & Abdelhalim, 2025). Moreover, AI not only supports students in completing complex tasks but also enables educators to provide richer, more responsive feedback (Guan, et al., 2025). The global shift toward digital education and the accelerated implementation of AI tools underscore the importance of teacher readiness, particularly among EFL pre-service teachers, who will soon be responsible for navigating this changing landscape.

The pedagogical potential of AI in language assessment lies in its capacity to expand traditional evaluation approaches. AI-driven assessment tools can enhance validity, provide real-time analytics, and foster formative assessment practices (Cope, et al., 2021). Teachers benefit from these innovations through increased efficiency and reduced subjectivity, while students gain access to more equitable and diverse forms of assessment. There is also a growing global demand for digital assessment, which propels the need for AI tools that align with both pedagogical and assessment goals (Cope et al., 2021). Nonetheless, while the affordances of AI are well-documented, its successful integration into assessment practices depends on educators' foundational understanding of assessment principles. This context demands a focus on the development of language assessment literacy (LAL) among EFL teachers, especially those in training, to ensure that AI is employed in ways that are both pedagogically sound and ethically responsible.

Despite the potential of AI in education, a core challenge remains which is the lack of sufficient LAL among pre-service teachers. Effective assessment requires more than technical know-how, it also necessitates a deep understanding of assessment constructs, validity, reliability, and fairness. In addition, the validity of AI-based assessments continues to be a major concern, especially when teachers lack the requisite knowledge to critically evaluate the tools they use (Xi, 2023). Therefore, English teachers must be adept not only in assessing learners' competencies but also in ensuring that the assessments they design are valid and reliable (Roslan, et al., 2022). Without a solid foundation in LAL, pre-service teachers may implement AI-based assessments in ways that are inconsistent with best practices, undermining both learning outcomes and assessment integrity.

In addressing these challenges, it is essential to examine the intersection between LAL and AI-based assessment readiness. The current educational landscape demands that pre-service teachers develop the competencies to integrate digital technologies without compromising assessment quality. Furthermore, LAL significantly influences teachers' ability to design and implement meaningful assessments (Zhang, et al., 2023). In tandem, research on AI-based assessment readiness reveals that teachers' perceptions and competencies with AI tools determine how effectively they can harness these technologies (Shahid, et al., 2024). Therefore, fostering LAL alongside readiness for AI integration

represents a strategic approach to equip pre-service teachers with the mindsets needed to navigate the digital transformation in language education.

Solutions to these challenges have begun to emerge in the form of targeted training and educational reforms. Reflective teaching practices can be embedded in teacher education to facilitate better decision-making regarding technology use (Xie, et al., 2019). This approach is particularly relevant in the context of AI, where educators must evaluate the suitability and ethical dimensions of technological tools. In addition, teacher training should include critical engagement with AI tools, allowing pre-service teachers to align AI capabilities with pedagogical and assessment objectives (Kasneci et al., 2023). Moreover, there is a need for structured guidance and support mechanisms that help educators make informed decisions about AI-based assessments (Viberg, et al., 2024).

Several previous studies also suggest for the integration of LAL into pre-service teacher education as a precursor to effective AI adoption. A comprehensive framework for LAL development should encompass technical skills, decision-making, pedagogical knowledge, and theoretical principles. When aligned with AI training, such a framework can equip pre-service teachers with the cognitive tools to critically evaluate the functionalities and limitations of AI-based assessments (Lan & Fan, 2019). In addition, noting that LAL provides the evaluative lens through which AI applications in assessment can be optimized (Swiecki et al., 2022). In the Indonesian context, professional learning community-based initiatives have demonstrated shifts toward assessment for/as learning and stronger use of authentic assessment (Saputra, et al,, 2020), while digital classroom implementations show gains in collaboration and engagement but also reveal uneven digital skills and access across pre-service cohorts (Rabbianty et al., 2025). Moreover, surveys in Indonesian public universities indicate that attitudes and intentions to adopt AI are shaped by performance and effort expectancy and are dampened by perceived risk (Helmiatin, et al., 2024). Consequently, the dual emphasis on LAL and AI readiness appears vital in preparing pre-service teachers to address the demands of modern language assessment.

Several empirical studies have sought to explore the individual constructs of LAL and AI readiness. Previous research examined the LAL of EFL teachers in Yemen and identified significant gaps, particularly in their understanding of assessment administration and underscored the importance of professional development programs focused on building LAL (Al-Akbari, et al., 2025). Another research investigated novice teachers' readiness for AI integration and found that limited training and familiarity with AI tools diminished their confidence and competence (Özer-Altınkaya & Yetkin, 2025). Meanwhile, in the AI-based assessment context, related research investigated psychological perspective, exploring how anxiety and resistance to change affect teachers' readiness for AI-based assessment (Shahid et al., 2024). Complementing these global patterns, Indonesian evidence from AI-assisted practicums shows that pre-service teachers benefit from idea generation and time efficiency but struggle with overreliance and ethical use—underscoring the need for mentored,

contextualized integration (Wulandari & Purnamaningwulan, 2024). At the institutional level, Indonesian adoption studies further highlight the role of perceived usefulness/ease and risk in shaping intentions to use AI (Helmiatin et al., 2024). These studies collectively indicate a pressing need for holistic frameworks that simultaneously address LAL and AI readiness. However, none of the existing literature directly examines the correlation between these two constructs in the context of EFL pre-service teachers.

Although studies on language assessment literacy and AI readiness have been conducted, their intersection remains under-researched, particularly in relation to how these two constructs shape pre-service teachers' capacity to integrate AI into assessment practices. This gap is especially relevant in the Indonesian context, where teacher education programs are still adapting to technological change and the incorporation of AI tools. Addressing this issue is the central contribution of the present study. This study aims to examine the connection between the language assessment literacy of EFL pre-service teachers and their readiness to use AI-based assessments in Indonesian universities. The participants were limited to pre-service teachers in Indonesian universities who had completed or were currently completing their teaching practice within one year to ensure that they had the most relevant and recent experience. To achieve its aim, this research explored two primary research questions: Is there any significant correlation between EFL pre-service teachers' language assessment literacy (LAL) and their readiness in using AI-based assessment? And How do EFL pre-service teachers perceive Language Assessment Literacy across its dimensions?

LITERATURE REVIEW

EFL Pre-Service Teachers and Their Professional Preparation

English as a Foreign Language (EFL) pre-service teachers are undergraduate students enrolled in teacher education programs preparing to become future English language educators. These pre-service teachers undergo pedagogical training and are expected to master essential aspects such as teaching competencies, subject knowledge, and language teaching methodologies, with an emphasis on pedagogical content knowledge (Yalcin Arslan, 2019). During the final stages of their academic journey, they typically engage in teaching practicum, allowing them to apply theoretical knowledge in real classroom settings. This phase is crucial as it facilitates the development of pedagogical skills, classroom management strategies, and assessment competence, all of which are critical in facing the diverse realities of classrooms (Brown, et al., 2021). In Indonesia, mentored practicums that incorporate AI-assisted lesson design streamline preparation and stimulate ideation but also surface risks of overreliance and ethical concerns, signaling the need for scaffolded supervision (Wulandari & Purnamaningwulan, 2024).

However, despite structured training, EFL pre-service teachers often encounter several instructional challenges. Among the most prominent issues are limited teaching

time, low student engagement, and poor student motivation (Taghizadeh & Ejtehadi, 2023). Another frequently reported issue is difficulty in implementing effective classroom assessments (Pardo, et al., 2024). These findings suggest the need for continuous professional support, innovation, and targeted development opportunities. Teacher education programs that provide robust, immersive professional experiences tend to better prepare pre-service teachers for the classroom (Lander et al., 2025). Moreover, structured programs that foster reflective practice significantly contribute to their professional readiness (Belford, et al., 2020). Evidence from Indonesian professional learning communities also indicates shifts toward assessment for/as learning and greater use of authentic tasks, suggesting locally viable routes to strengthen assessment competence (Saputra, et al., 2020).

Language Assessment Literacy among Pre-Service Teachers

Language Assessment Literacy (LAL) has become a fundamental component in teacher competence, especially for EFL pre-service teachers. It refers to teachers' knowledge, skills, and beliefs about designing, implementing, and interpreting language assessments (Lan & Fan, 2019). In addition, LAL entails the ability to ask critical questions concerning what to assess and how to assess it effectively (Arefian, 2022). LAL comprises both theoretical foundations and practical applications, positioning it as a core professional domain for language educators. Indonesian studies echo this view: collaborative teacher learning has been shown to foster enactment of authentic assessment while revealing uneven practical know-how that requires sustained support (Saputra et al., 2020).

There are four key dimensions within the LAL framework which are technical skills, score interpretation and decision-making, language pedagogy, and assessment theory and principles (Lan & Fan, 2019). A strong foundation in these dimensions equips teachers to align assessment tools with instructional objectives while addressing diverse learner needs. Research indicates that teachers with well-developed LAL are more capable of fostering student engagement and academic success through contextually appropriate assessment practices (Shafiee Rad, et al., 2024; Cui,et al., 2022).

Nevertheless, the literature also reveals significant gaps in LAL development. There are many educators, including pre-service teachers, are not adequately prepared to conduct effective assessments (Pastore & Andrade, 2019). In addition, assessment design remains one of the most difficult aspects for language teachers to master (Zhang et al., 2023). These findings highlight the urgent need for sustained professional development and assessment-focused training to enhance pre-service teachers' LAL. Similarly, Indonesian classroom implementations that use digital platforms for evidence of learning report improved engagement yet uneven digital skills, which may constrain assessment design quality if LAL is not explicitly developed (Rabbianty, et al., 2025).

Readiness for AI-Based Language Assessment

As educational technology evolves, artificial intelligence (AI) is increasingly integrated into assessment practices. AI-based assessment readiness refers to pre-service teachers' preparedness to incorporate AI tools effectively in classroom assessment. The readiness encompasses seven dimensions, which are performance expectancy, effort expectancy, social influence, facilitating conditions, perceived ease of use, perceived usefulness, and behavioral intention (Shahid et al., 2024).

In the digital era, integrating AI in assessment is increasingly seen as essential. Teachers must be able to utilize AI to streamline assessment procedures, provide immediate feedback, and enhance learning personalization (Ayanwale, et al., 2022). AI offers considerable benefits, such as reducing the time required to design and evaluate assessments while improving efficiency and student engagement (Xi, 2023; Swiecki et al., 2022). In Indonesian higher education, attitudes and intentions to adopt AI are shaped by performance and effort expectancy and are dampened by perceived risk, underscoring that readiness-building must address perceptions in addition to infrastructure (Helmiatin, et al., 2024).

Despite its advantages, the implementation of AI-based assessment presents new challenges. Aligning AI-generated content with instructional objectives remains a key difficulty for many pre-service teachers (Kohnke, 2025). Furthermore, ethical considerations, such as data privacy and assessment fairness, often go unaddressed (Celik, 2023). Field evidence from Indonesian AI-assisted practicums confirms these tensions—pre-service teachers gain efficiency and creativity but confront ethical use and quality-control concerns—highlighting the need for contextualized mentoring (Wulandari & Purnamaningwulan, 2024). These concerns underscore the need to equip pre-service teachers with both technical proficiency and critical awareness to ensure responsible use of AI in educational settings.

Interconnection Between Language Assessment Literacy and AI-Based Readiness

Understanding the interplay between LAL and AI-based assessment readiness is essential in modern teacher education. Although AI can facilitate assessment processes, its effective use depends on teachers' foundational knowledge of assessment principles. Educators often overlook core assessment concepts when integrating digital tools, especially AI (Arslan, 2025). Therefore, strong LAL is not merely complementary but foundational for meaningful and ethical AI implementation in assessments.

Teachers with high competence in both LAL and AI readiness are more likely to design assessments that are accurate, valid, and aligned with learning goals (Swiecki et al., 2022). This integrative approach enhances the quality of assessments and supports better learning outcomes (Zaim et al., 2024). Taken together, international findings emphasizing validity- and ethics-informed AI use align with Indonesian evidence: PLC-supported LAL growth

(Saputra et al., 2020) and UTAUT-based predictors of AI intention (Helmiatin et al., 2024) indicate complementary levers—professional learning for LAL and perception-shaping for AI adoption—while classroom studies also reveal local constraints (uneven digital skills; overreliance on AI) that may blunt benefits unless LAL explicitly mediates AI use (Rabbianty et al., 2025; Wulandari & Purnamaningwulan, 2024). While existing studies highlight important insights into LAL or AI readiness individually, the connection between these two domains remains largely unexplored. This study builds on that gap by considering them in tandem, especially among EFL pre-service teachers. Bridging this gap is essential to developing a future teaching workforce equipped for AI-enhanced educational landscapes.

Several empirical studies have explored related themes separately. The LAL of EFL teachers in Yemen was investigated and revealed that while teachers were confident in certain LAL dimensions, they reported deficiencies in assessment knowledge and expressed a need for professional training (Al-Akbari et al., 2025). Meanwhile, the AI readiness among Turkish pre-service teachers was found that despite recognizing the relevance of AI, many participants felt unprepared due to insufficient exposure and training opportunities (Özer-Altınkaya & Yetkin, 2025).

Additionally, in another study that investigated psychological factors influencing university lecturers' readiness for AI-based assessment in Malaysia found that anxiety negatively impacted adoption intentions, while resistance to change showed no significant effect (Shahid et al., 2024). Against this background, Indonesian literature illuminates *how* readiness and LAL can be advanced (via PLCs and practicum mentoring) but still stops short of testing *whether* LAL predicts AI readiness among pre-service EFL teachers—precisely the correlation this study addresses (Saputra et al., 2020; Wulandari & Purnamaningwulan, 2024; Helmiatin et al., 2024). However, none of these studies examined the direct correlation between LAL and AI-based assessment readiness in the context of pre-service EFL teachers. Thus, the current literature suggests a pressing need for integrated research examining how language assessment literacy may shape or predict AI-based assessment readiness. Investigating this connection could inform pre-service teacher education curricula and support more effective integration of assessment theory with technological innovation.

METHOD

Research Design

This study employed a mixed method explanatory sequential research design which means collected the quantitative data first, followed by qualitative data to explain the quantitative data (Creswell & Clark, 2018). This design was chosen because it allows statistical trends to be interpreted in light of participants' real experiences, thereby providing greater depth of understanding. In the first phase, a quantitative survey examined whether a significant correlation exists between EFL pre-service teachers' LAL and their

readiness to use AI-based assessment. In the second phase, qualitative interviews explored how participants with high and low levels of AI readiness differed in their perceptions of LAL, offering contextualized explanations for the patterns identified in the quantitative analysis. The combination of these two methods enhanced the validity of the study by connecting measurable outcomes with interpretive insights. This integration is essential because while correlation coefficients indicate the strength and direction of a relationship, they cannot reveal the underlying reasoning, beliefs, or pedagogical practices that explain the numbers. Therefore, the mixed-methods approach strengthened the credibility of the findings by linking generalizable patterns with contextualized explanations drawn from participants' lived experiences.

Population and Sample

For the quantitative phase, this research applied a purposive sampling to select participants based on particular criteria. The participants are EFL pre-service teachers who were currently enrolled in English Education programs and who had either completed or were currently undergoing their teaching practicum within one year prior to the data collection phase. Respondents were drawn from 60 various universities across Indonesia, thereby improving the heterogeneity and representativeness. Before participating, they were asked for their informed consent to ensure ethical consideration was upheld.

In the second phase, participants were drawn from those who had taken part in the quantitative stage and had expressed willingness to be further involved in the study. To gain deeper insight, extreme case purposive sampling was applied, targeting individuals whose responses were notably representative within the broader data (Cohen, et al., 2018). Therefore, it involved two participants that had the highest and lowest scores from the AI-based assessment of readiness since the initial findings highlighted this as a key area of interest. Although only two participants were included in the interview phase, this was a deliberate methodological choice consistent with the explanatory sequential design. Smaller samples may be sufficient when the study aim is focused, the sample is specific, the theoretical framework is established, and the data are rich (Malterud, et al., 2016). These conditions were fulfilled in the present study, as the aim centered on contrasting extreme cases of readiness, guided by the four dimensions of the LAL framework, with interviews generating detailed and meaningful insights. Thus, while small, the sample provided sufficient depth to contextualize and illustrate the quantitative findings. To address potential gender bias, both male and female participants were included. While contacting the sample for the interview, three expressed their reluctance. Therefore, another sample was invited.

Research Instruments

The primary data collection tool was a structured, close-ended questionnaire consisting of two major sections. The first section measured language assessment literacy,

and the second section assessed readiness for AI-based assessment. Both sets of items used a four-point Likert scale, representing four levels from strongly disagree to strongly agree, allowing respondents to select the extent to which they agreed with the statements given. The consideration to apply a four-point Likert scale is based on previous studies that suggest using an even scale in order to minimize respondents from consistently choosing neutral answers (Al-Akbari et al., 2025).

The questionnaire items for LAL adopted a previously developed questionnaire that consist four key dimensions which are technical skills, scores and decision-making, language pedagogy, and assessment theories and principles (Fitriyah, et al., 2022). These dimensions directly reflect the competencies that pre-service teachers are expected to develop during their training, thereby ensuring strong construct validity for this population. While the AI-based assessment readiness section was adopted based on the previous research which compromise seven dimensions that are performance expectancy, effort expectancy, social influence, facilitating conditions, users' anxiety, resistance to change, and attitude towards AI (Shahid et al., 2024). These dimensions are grounded in the UTAUT framework and psychological adoption models, both of which have been shown to predict technology integration in educational contexts. By selecting these validated and theoretically grounded instruments, the study ensured that the constructs measured were both reliable and relevant to the aim of exploring how assessment literacy shapes readiness to adopt AI-based assessment tools.

Before the distribution of the main questionnaire, the instrument underwent a pilot testing phase to ensure both validity and reliability. The pilot study involved 20 participants. The internal consistency of each section was measured using Cronbach's Alpha coefficient. The reliability coefficient for the LAL section was 0.891, and for the AI-based assessment readiness section, it was 0.876. Both values exceed the acceptable threshold of 0.70, indicating high reliability and internal consistency of the items used (Creswell & Creswell, 2022). Although the survey met the reliability criteria, several items were removed due to not meeting validity standards.

For content validity, the items were examined using the Pearson product-moment correlation, with a significance threshold set at $p < 0.05$. Out of the 18 LAL items, 15 met the significance criteria and were retained, while 3 items were excluded due to their failure to reach statistical validity. Similarly, out of the 27 AI readiness items, 19 were validated and retained for the final questionnaire. The remaining 8 items were removed because of either high insignificance or lack of alignment with the measurement objectives. Therefore, the questionnaire consists of 34 final total items.

To encourage participants to express their views freely and gain more information about Language Assessment Literacy, a qualitative approach was applied by collecting the data through interviews with selected representatives. These interviews were guided by a

modified version of the Language Assessment Literacy questionnaire, reformulated into open-ended questions translated into Bahasa Indonesia to collect deeper insights.

Data Collection Procedure

The data collection was conducted through online media using Google Forms, making the process efficient and accessible to participants from multiple regions. The online survey was distributed in two months via social media platforms commonly used by pre-service teachers, such as WhatsApp and Twitter, to ensure broader outreach and participation. Using Twitter, the researchers distributed the questionnaire trough several University Twitter base account that has an English Education study program. A total of 213 pre-service teachers voluntarily responded and completed the survey. However, 13 participants did not meet the criteria therefore only 200 participants (175 female, 25 male) consist of 192 undergraduate students and eight graduate students were included in the final analysis.

Participants were informed about the voluntary nature of their participation and the confidentiality of their responses. They were also given an approximate time estimate of 10–15 minutes to complete the questionnaire. No identifying information was collected, thus ensuring the anonymity of the participants in accordance with ethical research practices. The questionnaire consisted of three parts. The first part collected demographic data, the second part measured LAL, and the last part focused on AI-based assessment readiness. Each participant responded to a total of 34 validated items which consist of 15 LAL items and 19 for AI-based assessment readiness. All responses were automatically recorded and compiled via the Google Forms.

In the interview phase, all participants were informed about the study and asked for their consent prior to the interviews. The interviews were conducted online using WhatssApp as the media since the participants have no time because it was holiday at the moment this interview conducted. Furthermore, the instrument using LAL framework was translated to Bahasa Indonesia in order to make the participant easier in explaining the answer and to gain a deeper insight (Cirocki, et al., 2025).

Data Analysis

After the data were collected, they were exported to IBM SPSS version 27 for statistical analysis. The primary objective was to determine the correlation between EFL pre-service teachers' language assessment literacy and their AI-based assessment readiness. The responses to each questionnaire item were coded numerically based on the four-point Likert scale and with items 14 and 15 from the AI-based assessment readiness questionnaire reverse-coded due to present in negative statements. Composite scores were calculated for each respondent across both constructs. To test the correlation, Pearson's product-moment correlation coefficient was employed, as it is suitable for determining the

linear relationship between two continuous variables (Prestes, et al., 2021). Then, descriptive analysis was conducted to get clearer insight among each variable item.

For the qualitative data, after conducting interviews with participants, their answers were translated to English in order to make it coherence in language use, vocabulary, and idea delivery throughout the manuscript (Abfalter, et al., 2021). The interview data were then analyzed using thematic analysis (Braun & Clarke, 2006) with a hybrid inductive–deductive approach. This strategy has been recommended in mixed methods research because it balances openness to data-driven insights with theoretical alignment to established frameworks (Proudfoot, 2023). Following Braun and Clarke's six steps, the analysis began with familiarization and inductive coding of participants' responses using Atlas.ti 24. These codes were subsequently organized into broader themes. In the final stage, the themes were deductively examined in relation to the four dimensions of the Language Assessment Literacy (LAL) framework. This process ensured that the analysis was both grounded in participants' experiences and theoretically aligned with the LAL framework. To enhance credibility, member checking was conducted by sharing preliminary interpretations with selected participants to confirm the accuracy of the findings with their perspectives.

RESULTS AND DISCUSSION

Results

The Relation between Language Assessment Literacy and AI-based Assessment Readiness

This section provide the findings derived from participant data, starting with aanalyzing the correlation between Language Assessment Literacy and AI-based Assessment Readiness, which offer insights into the data's central tendencies and variability. Then followed by descriptive statistical summary of the variables to get more deeper insight. These findings reveal the nature, strength, and direction of the relationships between the variables. After collecting the data through a close-ended questionnaire distributed to 200 qualified participants from 60 different universities in Indonesia, the responses were analyzed using SPSS version 27.

Table 1.
Pearson Correlation Result

		Language Assessment Literacy	**AI-based Assessment Readiness**
Language Assessment Literacy	Pearson Correlation	1	0.494**
	Sig. (2-tailed)		0.000
	N	200	200
AI-based Assessment	Pearson Correlation	0.494**	1
	Sig. (2-tailed)	0.000	

Readiness	N	200	200

Table 2.
Pearson Correlation Interpretation

The Correlation Coefficient	Interpretation
0.20-0.35	Low correlation
0.35-0.65	Moderate correlation
0.65-0.85	Strong correlation
>0.85	Very strong correlation

To determine the relationship between LAL and AI-based assessment readiness, Pearson Product-Moment Correlation was conducted. Based on Table 1, the correlation was significant between LAL and AI-based assessment readiness, with a significant value of .000. While the correlation value is .494, which means they have a medium correlation based on correlation interpretation as seen in Table 2 (Cohen et al., 2018). Based on Figure 1, the scatter plot graphic shows a positive relationship between the two variables, as indicated by the dots near the linear line being close to each other. This suggests that higher Language Assessment Literacy is associated with greater readiness to use AI-based assessment.

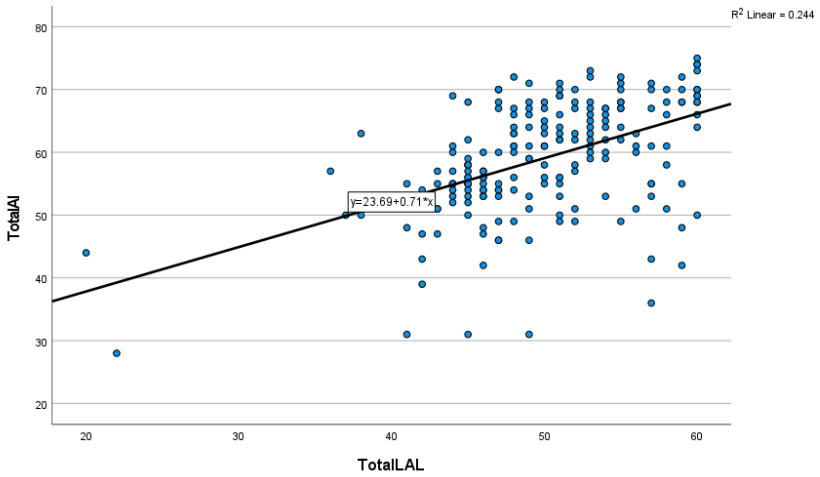

Figure 1. Scatter plot graphic of LAL and AI-based Assessment Readiness

Table 3.
4-Point Scale Mean Range with Four Levels

Mean Range	Interpretation
1.00 -1.74	Low
1.75 -2.49	Moderate Low
2.50 -3.24	Moderate-High
3.25 -4.00	High

Table 4.
EFL Pre-service Teachers Language Assessment Literacy (Statement 1-15)

Dimension	Statement	Mean	Interpretation
Technical skills	I have sufficient knowledge using ready-made test from textbook packages or from other sources.	3.46	High
	I understand how to give feedback to students based on information from tests/assessment.	3.45	High
	I am aware of how to use self- or peer-assessment.	3.33	High
	I am knowledgeable about using informal, continuous, non-test type of assessment is well formed in my mind.	3.39	High
	I am well-versed in using the Language Portfolio, an adaptation of it or some other portfolio.	3.08	Moderate High
Scores and decision making	I am well-informed about giving grades.	3.41	High
	I have sufficient knowledge how to find out what needs to be taught/learned.	3.38	High
Language Pedagogy	I am well-versed in testing/Assessing Receptive skills (reading/listening).	3.38	High
	I have enough knowledge in testing/Assessing Productive skills (speaking/writing).	3.31	High
	Testing/Assessing microlinguistic aspects (grammar/vocabulary) is well formed in my mind.	3.16	Moderate High
	I am knowledgeable about testing/Assessing Integrated language skills.	3.24	Moderate High
	I am knowledgeable about testing/Assessing aspects of culture (students' attitudes).	3.49	High
Theories and principles	I am aware of how to establishing reliability of tests/assessment.	3.42	High
	I am well-versed in establishing validity of tests/assessment.	3.29	High
	I understand how to use statistics to study the quality of tests/assessment.	3.21	Moderate High
	Total	3.33	High

Furthermore, the researchers used descriptive quantitative analysis to analyze the Language Assessment Literacy (LAL) and AI-based Assessment quantitative data in order to gain deeper and interpret data in terms of means and interpret data of each variable based the 4-Point Scale Mean Range with Four Level as seen in Table 3. This finding underscore the strength level and direction of the relationship between variables. The findings for the LAL variable show an overall average score of 3.33 among all participants. This suggests that EFL pre-service teachers generally hold high levels LAL with 11 items are considered high and only four items that are rated moderate high.

A trend also reflected across each individual dimension. As shown in Table 4, the Score and Decision Making dimension is the only dimension which all items are rated high, while technicall skills, language pedagogy, and theories and principles consist about one up to two items with moderate high. The item with the lowest mean was the use of language portofolio with a mean of 3.08. Another notable point is that the highest mean from LAL was related to understanding in testing cultural aspect with a mean score of 3.49 which highlight

their knowledge in accordance with language pedagogy. Overall, these findings reflect a strong foundation in LAL among EFL pre-service teachers.

Table 5.
EFL Pre-service Teachers AI-based Assessment Readiness (Statement 16-34)

Dimension	Statement	Mean	Interpretation
Performance Expectancy	Using the AI-based assessment system can enhance my productivity.	3.39	High
	Using the AI-based assessment system can enhance the learning performance of students.	3.12	Moderate High
	Using an AI-based assessment system improves the accuracy of students' assessment.	3.08	Moderate High
	Using the AI-based assessment system would help to reduce my workload.	3.45	High
Effort Expectancy	It would be easy for me to become skillful at using AI-based assessment systems.	3.20	Moderate High
	The use of AI-based assessment system can be easy for me.	3.38	High
	Learning the use of AI-based assessment system can be easy for me.	3.34	High
	My interaction with AI-based assessment system can be clear and understandable.	3.21	Moderate High
Facilitating Conditions	I have necessary resources to use AI-based assessment system.	3.21	Moderate High
	The AI-based assessment system would be compatible with other technologies or systems I use.	3.21	Moderate High
Social Influence	My colleagues support the use of the AI-based assessment system.	3.21	Moderate High
	My superiors would support the use of the AI-based assessment system.	3.19	Moderate High
	The university's top management would support the use of the AI-based assessment system.	3.01	Moderate High
Users' Anxiety	It scares me to think that it can cause AI-based assessment system to destroy a large amount of information by hitting the wrong key. R[1]	2.00	Moderate Low
Resistance to Change	I do not agree with the changes entailed by the use of AI for the assessment. R[1]	2.39	Moderate Low
Attitude towards AI	AI-based assessment system is an appropriate tool for me to use	3.09	Moderate High
	I like the idea of using the AI-based assessment system for assessing students' performance	3.19	Moderate High
	I think using the AI-based assessment system will be advantageous for assessing students' performance	3.25	High
	Overall, my attitude toward using the AI-based assessment system is positive	3.28	High
	Total	3.11	Moderate High

[1]Note. The scoring of items marked with R was reversed

The next part of the findings focus at how ready pre-service teachers in adapting AI-based assessment, which produced an overall average score of 3.11, indicating a generally

moderate high readiness level. However, variations emerged among the different dimensions of this variable. The highest-scoring dimension was Effort Expectancy, particularly on statements indicating that learning and interacting with AI-based systems was perceived as relatively easy, with means ranging from 3.20 to 3.38. Similarly, Performance Expectancy received high ratings, with the highest item score is 3.45 showing that participants believed AI could help reduce their workload.

On the other hand, Users' Anxiety and Resistance to Change were the lowest rated dimensions, with means of 2.00 and 2.39 respectively. These scores reflect that some participants still feel hesitant about adopting AI-based tools in assessment. Other dimensions such as Facilitating Conditions, Social Influence, and Attitude Towards AI were consistently rated as moderate high, suggesting a generally positive but cautious outlook. These findings indicate that while EFL pre-service teachers are relatively open about using AI in assessment, in some areas especially related to emotional readiness and institutional support, may require further attention and development. Moreover, the descriptive analysis between LAL and AI-based assessment provide more information about the interplay direction among them which is moderate high. This result served as the foundation for conducting the qualitative phase in this mixed methods research.

EFL Pre-service Teachers' Perception of Language Assessment Literacy

This section reports findings from interviews with two EFL pre-service teachers, Adi and Ria (pseudonyms), who were selected through extreme case purposive sampling. One participant represented a high level of AI-based assessment readiness, while the other reflected a low level. Both voluntarily agree to participate in the study. This qualitative phase aimed to explore how their perspectives on Language Assessment Literacy may differ, providing deeper insight into the contrast identified in the quantitative results.

Technical Skills

Both pre-service teachers showed fundamental knowledge of how to assess textbook-based tests. They demonstrated an understanding of important concepts, including validity as well as reliability. However, the depth of their responses varied significantly. Nevertheless, the level to which they elaborated on this differed greatly. Adi elaborated on the fact that test quality must match the learning goals and be flexible based on test difficulty as well as the needs of students. Adi remarked the following:

> "If the test is too easy, too difficult, or not aligned with the material, it is better to adapt it before using it."

This difference implies that higher AI readiness may support a more reflective and adaptive approach to test evaluation.

The high score response demonstrates his skills to make instructional decisions that consider content validity, task variation, and the learner context. On the other hand, Ria provided a more surface-level response, with just three criteria being stated as validity, reliability, and authenticity. There was no elaboration and application from her response. This difference implies that higher AI readiness may support a more reflective and adaptive approach to test evaluation.

In discussing the role of feedback in language assessment, a distinct difference emerged between the two participants' responses, particularly in terms of depth, delivery, and pedagogical orientation. The high AI-readiness participant, emphasized a motivational and student-centered perspective on feedback. He described feedback as more than just a corrective tool, suggesting that it should be thoughtfully delivered in ways that promote positive and reflective approaches. Adi expressed it as:

> "Feedback should be delivered using positive and motivating language… involve students to discuss their results so they are more reflective."

His statement reflects a clear understanding that feedback is not only pointing out errors, but also about facilitating response between teacher and student that encourages learners to take ownership of their progress. This shows his grasp of feedback as both an instructional strategy and a developmental tool that helps learners engage in self-regulation and improvement over time. In contrast, the low score' answer was more corrective. Ria stated that:

> "The feedback is expected to help students understand their weaknesses and improve their abilities in the future."

While this response acknowledges one of the core functions of feedback, it lacks elaboration on how feedback should be communicated or how students are expected to engage with it. The contrast highlights the broader pedagogical implications embraced by participants with higher readiness. Furthermore, when discussing alternative assessments, both acknowledged the use of peer and self-assessment, but with differing levels of pedagogical depth. Adi, the participant with high AI-based assessment readiness, offered a well-thought-out and process-oriented perspective. He pointed out the following:

> "To implement alternative assessments such as self-assessment or peer assessment in the classroom, teachers can start by providing a clear scoring rubric, then train students on how to use it through examples… followed by a brief reflection on what can be improved."

This demonstrates an understanding of assessment literacy development among learners. The mention of training students using examples and incorporating reflection illustrates his awareness that students may not intuitively understand how to assess

effectively without guidance and modeling. On the other hand, Ria's explanation of alternative assessment was more concise:

> "[In self assessment] by providing students with a scoring rubric for a specific task, along with reflection questions. [While peer assessment] students assess their peers' work and provide constructive feedback."

While her answer covered the general idea, it lacked elaboration on how to implement these practices effectively. This contrast suggests that while both participants recognize the role of alternative assessments, only the high-readiness participant demonstrated a more comprehensive view. The gap reflects how AI-readiness might also influence one's capacity to critically consider instructional design in assessment.

Similarly, participants' perceptions of informal assessment and portfolio use underscored meaningful differences in how assessment practices are conceptualized. Adi described informal assessment as an ongoing learning tool involving the integration of daily classroom interactions such as observations, reflections, and teacher-student dialogues as valuable sources of assessment data. He also remarked "The portfolio not only presents final outcomes but also highlights the learning process." His answer indicates that he perceives assessment as integral to formative development. In contrast, the participant with low AI-based assessment readiness expressed a more functional understanding of informal assessment practices. Ria provided the following insight:

> "Observation, self-assessment, task-based, portfolio... we can use the portfolio to identify students' progress and provide feedback."

While she correctly identified basic informal assessment tools, her explanation focused more on the product of learning rather than the process. These differences reveal how higher AI-based readiness may foster deeper pedagogical integration of assessment practices.

Score and Decision Making

In the aspect of score interpretation and decision-making, both participants acknowledged the use of incorporating multiple assessment components into final grade calculations. However, the depth and clarity of their responses reflected varying levels of pedagogical understanding. Adi shared his view:

> "… to combine various assessment results can be done by setting score weights to each type of assessment. For example, tests could be worth 40%, homework 20%, projects 25%, and class participation 15%, calculates them according to their weights, and sums them up to get the final grade. It is also important to ensure that all assessments reflect the learning objectives and suit with the students' abilities, so the final grade is fair and accurate."

This response indicates a well-informed awareness of the need to not only distribute grading weights across multiple components but also to ensure that each component aligns with clearly defined learning outcomes. His answer suggests that he also views grading as a reflective practice grounded in fairness, transparency, and alignment with educational purposes. In contrast, Ria emphasized *"Combine results from tests, assignments, and participation based on weights."* Her answer demonstrates basic awareness of assessment procedures, but lacking explanation of how those weights are determined or used to ensure fairness. The difference suggests that higher AI-based assessment readiness supports a more reflective and principled related to grading.

Moreover, in discussing the role of assessment results in instructional decision-making, both participants acknowledged the importance of using student performance data to inform teaching. However, their responses differed significantly in terms of depth and pedagogical orientation. Adi highlighted this perspective *"... use assessment results to identify students' learning needs is by analyzing their answer patterns or performance, then identifying which areas they have not yet mastered."* His response demonstrates an understanding of how assessment can be used not just to evaluate students, but to guide differentiated instruction based on specific learning gaps. Ria in contrast, while offering a similar view, remained more focused on correction than instructional design *"From assessment results, we can analyze weaknesses and design materials to help them develop better."* Her answers focus remains on identifying weaknesses and producing supporting materials. This reinforces the pattern where the high-readiness participant applies assessment for forward planning, while the low-readiness participant limits its use to remediation.

Language Pedagogy

When asked how receptive skills should be assessed, both participants were able to identify relevant task types, but their responses varied considerably in terms of depth, purpose, and pedagogical awareness. In testing receptive skills, the participant with high score identified the need to assess cognitive skills through comprehension tasks. Adi pointed out the following:

> "In my opinion, an appropriate method to assess receptive skills like reading or listening is by providing a text or audio followed by comprehension questions which measure the ability to get information, main ideas, and details."

This response shows his understanding of testing depth of comprehension beyond recall. Meanwhile, the low score participant responded by listing a set of task formats without elaboration. She mentioned several type of test such as multiple choice, scanning, skimming, summarizing, and cloze task. The lack of conceptual connection to cognitive demands suggests a more superficial understanding of skill assessment. This difference is

important because it highlights how higher levels of AI-readiness may correlate with a more reflective and pedagogically informed approach to assessment.

In assessing productive skills, a noticeable difference emerged in how each participant approached the evaluation of students' speaking and writing abilities. Adi suggested authentic, rubric-based tasks such as essays and presentations. His response revealed the following perception *"In my opinion, the appropriate method to assess productive skills use presentations or essays with rubrics that assess content, structure, vocabulary, and fluency."* This answer points to his awareness of the needs for comprehensive, fair, and transparent criteria covering different aspects of language performance. In contradistinction, Ria focused more on the traditional methods such as read aloud, short writing, error recognition, and essays. Her answer indicating limited attention to task design or assessment criteria. This highlights how higher AI-readiness may align with a deeper understanding of productive skill evaluation.

Moreover, when discussing the assessment of grammar and vocabulary, the two participants demonstrated a clear differences in both conceptual depth and practical application. The high AI-readiness participant proposed integration into meaningful language tasks. Adi shared his view:

> "The appropriate method to assess grammar and vocabulary is through written tests such as fill-in-the-blanks, multiple choice, or sentence correction. It can also be done through writing or speaking tasks that are analyzed using a rubric focusing on the accurate and varied use of structures and vocabulary."

His responses implied a communicative approach in assessing grammar and vocabulary. In contrast, the low AI-readiness participant, presented a more traditional and test-oriented view of grammar and vocabulary assessment such as multiple choice, gap filling, and cloze task. Her answer reflecting a test-focused mindset. The contrast illustrates how high AI-readiness correlates with contextualized language assessment practices. Regarding integrated skills and culture, a clear distinction emerged between the high and low AI-readiness participants. The high score participant proposed project-based learning that reflects real-life communicative demands. Adi shared his view:

> "In my opinion, a method that can be used to assess all language skills at the same time is project-based tasks, such as presentations, debates, or making videos, which integrate listening, reading, writing, and speaking skills in one integrated activity."

This response reflects not only a practical approach but also a sophisticated grasp of integrated skill assessment. In addition to his emphasis on skills integration, he also provided thoughtful engagement with cultural aspect competence. He remarked the following:

> "… assessing cultural aspects in language learning is necessary, because language and culture are related to each other. Cultural understanding helps students use the language more appropriately and contextually. The assessment can focus on understanding and comparing cultural differences critically and openly."

His answer signals more than a surface-level appreciation for culture. It also indicates an awareness of the pedagogical value of cultural content in fostering critical thinking. On the other hand, regarding assesing integrated skills, Ria's response consisting of a list of task formats. She named several test types, including task-based, project-based, presentation. Her explanation lacked pedagogical specificity, which may indicate a less developed understanding of how assessment tasks function in relation to language proficiency and communicative demands. Furthermore, related to assessing cultural aspect, she also agreed with the importance of culture but offered a simpler perspective:

> "I think it is necessary, because culture and language are related to each other. Language learning will be easier and more meaningful when cultural elements are included, especially if it is relevant with the context of the students."

While her statement affirms the role of culture in enriching the learning experience, it lacks the analytical and pedagogical elaboration found in Adi's response. Her comment emphasizes motivation and relevance, which are important, yet it does not delve into how culture can be critically assessed. This contrast underscores how AI-readiness may influence how deeply culture and integration are embedded in assessment design.

Theories and Principles

On the concepts of reliability and validity, both participants gave correct definitions, but the depth of their understanding differed noticeably. The high readiness participant, demonstrated a clear grasp of both concepts and their practical classroom implications. Adi defined reliability as a test's ability to produce consistent results and explained that it can be improved by ensuring clear questions, level-appropriate difficulty, sufficient item quantity, and consistent test procedures. He elaborated:

> "Teachers can make sure that the questions are clear, appropriate for the students' level, the number of questions is sufficient, and the test procedure is carried out consistently without bias."

This shows a pedagogically grounded awareness that reliability also depends on fair test administration. In discussing validity, Adi again emphasized instructional alignment *"Design questions that are relevant to the material, appropriate to the students' level, and cover the targeted competencies."* His discussion of content relevance and competency coverage suggests greater sophistication about the role of assessment to support learning objectives and instruction. Ria, the participant with the lowest AI-based readiness, showed similar understanding of the definitions, but her responses were briefer and contained less

thorough instruction-based reflection. She defined reliability as when the test always gives the same results.

Similarly, on the definition of validity, she stated *"[Validity is] a test that measures what it is supposed to measure. One way to improve it is by designing questions that align with the learning objectives and the material."* Her response demonstrates awareness of basic content alignment, but without deeper insight into how tasks reflect competencies or student levels. Overall, Adi's explanations reflect a more developed understanding of assessment principles, connecting them directly to instructional quality and fairness. Ria's answers, though accurate, remain at a surface level. This contrast supports the pattern found in earlier sections, reinforcing that higher AI-readiness tends to align with deeper, more pedagogically informed views of Language Assessment Literacy.

Finally, in discussing the role of statistics in assessment, both participants recognized its relevance, yet their depth of understanding diverged significantly. Adi expressed a well-rounded view:

> "... statistics help teachers understand how well students have mastered the material, observe score distribution, and identify items that are too easy or too difficult [in order to] help in designing further lessons."

This response reflects his awareness of the diagnostic and formative power of statistical data. His mention of score distribution and item difficulty shows a familiarity not just with the numbers themselves, but with how these numbers translate into actionable insights for lesson adjustment, student support, and future assessment design. Ria's explanation was more basic *"The role of statistics is to assist in analyzing test results to understand students' level of comprehension and identify their needs."*

The low participant lacking detail on how statistics can be used diagnostically. While her statement is not incorrect, it lacks the elaboration necessary to demonstrate deep engagement with the concept. Her response implies a general understanding that data can inform instruction. This contrast illustrates that deeper assessment literacy may be connected to higher AI-readiness levels, particularly in interpreting data for pedagogical improvement.

Discussion

In this study, the researchers chose not to examine the results from each method on their own. Instead, the design of the research allowed the qualitative insights to add depth and meaning to the quantitative findings. By bringing the two approaches together, the researchers were able to build a more complete and grounded interpretation of the data. The findings reported there is significant correlation between language assessment literacy and AI-based assessment readiness among EFL pre-service teachers. It means that the more literate pre-service teachers are towards language assessment, the more ready they are to

apply AI-based assessment for the learning process. This result supports the conceptual argument that integrating technology into education is not only about technical competency but also is fundamentally associated with pedagogic knowledge (Arslan, 2025). It further implies that when pre-service teachers have a stronger foundation in assessment principles, they are more likely to feel prepared to adopt and use AI tools for assessment.

The finding further presents that the degree of relation between the two variables is a positive moderate correlation. It means that although the two constructs are clearly correlated, various other factors can also contribute to the readiness of teachers, such as personal pedagogic knowledge, cultural aspects, or support from the institution, which most probably are playing important roles (Celik, 2023; Zaim et al., 2024). Furthermore, it underscores the idea that effective technology adoption requires a holistic approach, where foundational pedagogical competence acts as a critical consideration for navigating new technological landscapes (Kohnke, 2025). Therefore, as the moderate positive correlation confirms the significant association between language assessment literacy and readiness for assessment based on AI, it also highlights the necessity of a broader approach that integrates pedagogical and educational institutions to fully support effective AI integration in assessment.

Furthermore, the quantitative findings showed an overall high level of LAL (M=3.33) in EFL pre-service teachers. This is similar to the earlier research that discovered high levels of LAL in novice teachers (Fitriyah et al., 2022). In contrast with some earlier research that showed extensive gaps in teachers' assessment (Al-Akbari et al., 2025; Pastore & Andrade, 2019). Nonetheless, the qualitative findings present a more diverse perspective. Even though the two participants were able to define the essential concepts, the participant with a high level of AI readiness showed constantly deeper and more pedagogically based knowledge. This suggests that high LAL scores on a survey may provide underlying differences in the ability to apply principles in practice, a distinction that is critical for effective teaching (Al-Akbari et al., 2025; Pastore & Andrade, 2019). Consequently, although the overall high level of LAL is promising, the qualitative data presents the true pedagogic impact as lying on the understanding not only on the conceptual level but also on the depth on the practical level to support the greater attention on teacher education programs to build up not just the conceptual knowledge but also the practice-based competency.

Additionally, the finding that shows pre-service teachers have a "moderate high" readiness for AI-based assessment (M=3.11) aligns with a previous study on teacher perceptions of AI. The high scores in dimensions like performance expectancy and effort expectancy reflect a general optimism about AI's potential to enhance efficiency (Ayanwale et al., 2022; Swiecki et al., 2022). Nevertheless, the relatively low level of users' anxiety and resistance to change is also found by other research where teachers are shown to fear their lack of technical skills and the disruption that might occur due to the utilization of AI (Shahid et al., 2024; Özer-Altınkaya & Yetkin, 2025). This difference of apprehension suggests that

readiness is not a simple state of acceptance but a complex acceptance between perceived benefits and personal or systemic barriers (Viberg et al., 2024).

Moreover, the qualitative interviewing of one high score and one low AI-based assessment readiness participant served to give insightful rationale for the quantitative data. The high score responded with a deep, principled knowledge of LAL, relating assessment practices to pedagogy, fairness, student motivation, and reflective practice, while the low score provided for a surface-level response. The top-score ability to critically think about why and how to assess is exactly the type of higher-order thinking required to assess and utilize the complex AI tools ethically (Kasneci et al., 2023). It appears that being truly ready to utilize AI in assessment is not only based on simple knowledge but also on deep and reflective knowledge about language assessment principles, which helps in thinking critically and adapting effectively (Lan & Fan, 2019; Cope et al., 2021; Guan et al., 2025). Overall, the qualitative results reaffirm that actual preparation for the use of AI-based assessment requires more than surface-level familiarity. It also needs deep and thoughtful knowledge of language assessment literacy to support pre-service teachers in critically reflecting, modifying, and ethically incorporating AI into their teaching practice.

In addition, the findings hold significant importance for EFL teacher education programs. Professional preparation on the use of AI tools should not just be technology workshop-focused but must be considerably integrated with the principles of language assessment. Curricula must be redesigned to build more integrated learning experiences. Teacher education programs can effectively prepare pre-service teachers to take advantage of AI when they provide experiences, encourage ethical knowledge, and support continuing professional learning. Institutions are expected to fulfil these needs to prepare for assisting the meaningful integration of AI across education (Kohnke, 2025). This includes providing pre-service teachers with the capabilities to carefully choose, adapt, and ethically control the use of AI tools to support the pedagogical goals, rather than allowing the technology to dictate the pedagogy (Belford et al., 2020; Pardo et al., 2024; Yalcin Arslan, 2019).

The results also present implications for action at the institution and policy levels. The facilitating conditions, users' anxiety, and resistance to change for the AI readiness dimension directly indicate the need to afford institutional support. Ministries of education and universities need to look beyond advocacy for the use of AI and give real resources. That means providing access to safe technology and software for AI, as well as providing ongoing continuing professional development, along with clear guidelines for the use of AI for assessment purposes (Kohnke, 2025). Unless there is support of this kind, the good intentions of pre-service teachers behind the use of AI will not change into good practice when they encounter the realities of the classroom (Lander et al., 2025; Roslan et al., 2022). Therefore, advancing AI-based assessment readiness requires not only curricular reform but also systemic institutional and policy-level support to ensure that pre-service teachers' positive intentions are shown in effective classroom practices. Therefore, advancing AI-

based assessment readiness requires not only curricular reform but also systemic institutional and policy-level support to ensure that pre-service teachers' positive intentions are shown into effective classroom practices.

CONCLUSION

The significant positive relationship between the language assessment literacy and AI-based assessment readiness suggests that if an individual has a higher LAL, that person will have a readiness to use AI for assessment. This study has practical contributions to English language education, language assessment literacy, and AI readiness by expanding current understandings and providing a foundation for future research. Practically, the implications of this study suggest that teacher education programs should be responsible for two embedded complex elements within their programs in order to make the pre-service teachers not only be aware of assessment principles but also be able to use those principles in evolving technological contexts. Therefore, curriculum development should consider embedding reflective, practical components that support the responsible and effective use of AI in assessment contexts.

From the results and limitations, there are a number of possibilities for future research. A longitudinal study could follow pre-service teachers as they transition into their professional careers to see how their LAL and AI-based assessment readiness change as they gain real-life teaching experience. This study could be carried out with in-service teachers, or in a different national context to see to what extent the level of the correlation fits. Future research could also examine other potential aspects of AI-based assessment readiness, such as teachers' epistemic beliefs or even personality traits.

ACKNOWLEDGEMENT

This study was conducted under the Structural Research Program (*Program Penelitian Struktural*) and supported through research funding granted by the Faculty of Cultural Studies, Universitas Brawijaya, under SK Dekan No. 59/2025. The authors sincerely appreciate this financial support.

REFERENCES

Abfalter, D., Mueller-Seeger, J., & Raich, M. (2021). Translation decisions in qualitative research: A systematic framework. *International Journal of Social Research Methodology*, *24*(4), 469–486. https://doi.org/10.1080/13645579.2020.1805549

Al-Akbari, S., Nikolov, M., & Hódi, Á. (2025). EFL teachers' language assessment literacy training needs. *Social Sciences & Humanities Open*, *11*, 101254. https://doi.org/10.1016/j.ssaho.2024.101254

Arefian, M. H. (2022). Perceptions of self-assessment literacy and self-directed reflection during online learning for Iranian EFL student teachers. *Reflective Practice*, *23*(6), 623–634. https://doi.org/10.1080/14623943.2022.2096584

Arslan, S. (2025). English-as-a-foreign language university instructors' perceptions of integrating artificial intelligence: A Turkish perspective. *System*, *131*, 103680. https://doi.org/10.1016/j.system.2025.103680

Ayanwale, M. A., Sanusi, I. T., Adelana, O. P., Aruleba, K. D., & Oyelere, S. S. (2022). Teachers' readiness and intention to teach artificial intelligence in schools. *Computers and Education: Artificial Intelligence*, *3*, 100099. https://doi.org/10.1016/j.caeai.2022.100099

Belford, N., Tudball, L., & Kewalramani, S. (2020). Research-informed 'immersive' professional learning with pre-service teachers in schools. *Cogent Education*, *7*(1), 1733223. https://doi.org/10.1080/2331186X.2020.1733223

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brown, A. L., Myers, J., & Collins, D. (2021). How pre-service teachers' sense of teaching efficacy and preparedness to teach impact performance during student teaching. *Educational Studies*, *47*(1), 38–58. https://doi.org/10.1080/03055698.2019.1651696

Celik, I. (2023). Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. *Computers in Human Behavior*.

Cirocki, A., Anam, S., Drajati, N., & Soden, B. (2025). Assessment Literacy Among Indonesian Pre-service English Language Teachers: A Mixed-Methods Study. *Iranian Journal of Language Teaching Research*, *13*(1). https://doi.org/10.30466/ijltr.2025.55691.2851

Cohen, L., Manion, L., & Morrison, K. (2018). *Research Methods in Education* (Eight edition). Routledge.

Cope, B., Kalantzis, M., & Searsmith, D. (2021). Artificial intelligence for education: Knowledge and its assessment in AI-enabled learning ecologies. *Educational Philosophy and Theory*, *53*(12), 1229–1245. https://doi.org/10.1080/00131857.2020.1728732

Creswell, J. W., & Clark, Vicki. L. P. (2018). *Designing and Conducting Mixed Methods Research* (Third edition). Los Angeles: SAGE.

Creswell, J. W., & Creswell, J. D. (2022). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (6th ed.). SAGE Publications, Inc.

Cui, Y., Liu, Y., Yu, H., & Gao, Y. (2022). Developing English teachers' language assessment literacy in an EAP reform context through test design: A case study. *System*, *109*, 102866. https://doi.org/10.1016/j.system.2022.102866

Fitriyah, I., Masitoh, F., & Widiati, U. (2022). Classroom-based language assessment literacy and professional development need between novice and experienced EFL teachers. *Indonesian Journal of Applied Linguistics*, *12*(1), 124–134. https://doi.org/10.17509/ijal.v12i1.46539

Guan, L., Lee, J. C.-K., Zhang, Y. E., & Gu, M. M. (2025). Investigating the tripartite interaction among teachers, students, and generative AI in EFL education: A mixed-methods study. *Computers and Education: Artificial Intelligence*, 100384. https://doi.org/10.1016/j.caeai.2025.100384

Helmiatin, Hidayat, A., & Kahar, M. R. (2024). Investigating the adoption of AI in higher education: A study of public universities in Indonesia. *Cogent Education, 11*(1), 2380175. https://doi.org/10.1080/2331186X.2024.2380175

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kohnke, L. (2025). Preparing future educators for AI-enhanced classrooms: Insights into AI literacy and integration. *Computers and Education: Artificial Intelligence*, *8*, 100398. https://doi.org/10.1016/j.caeai.2025.100398

Lan, C., & Fan, S. (2019). Developing classroom-based language assessment literacy for in-service EFL teachers: The gaps. *Studies in Educational Evaluation*, *61*, 112–122. https://doi.org/10.1016/j.stueduc.2019.03.003

Lander, N., Mazzoli, E., Orr, J., Verswijveren, S. J. J. M., Contardo Ayala, A. M., Brown, T., & Muir, T. (2025). TransformEd: An innovative model to support professional growth of pre-service teachers. *Teaching Education*, 1–22. https://doi.org/10.1080/10476210.2025.2462312

Malterud, K., Siersma, V. D., & Guassora, A. D. (2016). Sample Size in Qualitative Interview Studies: Guided by Information Power. *Qualitative Health Research*, *26*(13), 1753–1760. https://doi.org/10.1177/1049732315617444

Mekheimer, M., & Abdelhalim, W. M. (2025). The digital age students: Exploring leadership, freedom, and ethical online behavior: A quantitative study. *Social Sciences & Humanities Open*, *11*, 101325. https://doi.org/10.1016/j.ssaho.2025.101325

Özer-Altınkaya, Z., & Yetkin, R. (2025). Exploring pre-service English language teachers' readiness for AI-integrated language instruction. *Pedagogies: An International Journal*, *0*(0), 1–17. https://doi.org/10.1080/1554480X.2025.2451299

Pardo, R., García-Pérez, D., & Panadero, E. (2024). Shaping the assessors of tomorrow: How practicum experiences develop assessment literacy in secondary education pre-service teachers. *Teaching and Teacher Education*, *152*, 104798. https://doi.org/10.1016/j.tate.2024.104798

Pastore, S., & Andrade, H. L. (2019). Teacher assessment literacy: A three-dimensional model. *Teaching and Teacher Education*, *84*, 128–138. https://doi.org/10.1016/j.tate.2019.05.003

Prestes, P. A. N., Silva, T. E. V., & Barroso, G. C. (2021). Correlation analysis using teaching and learning analytics. *Heliyon*, *7*(11), e08435. https://doi.org/10.1016/j.heliyon.2021.e08435

Proudfoot, K. (2023). Inductive/Deductive Hybrid Thematic Analysis in Mixed Methods Research. *Journal of Mixed Methods Research*, *17*(3), 308–326. https://doi.org/10.1177/15586898221126816

Rabbianty, E. N., Virdyna, N. K., Azizah, S., Roda'i, M., & Shidqi, Z. A. (2025). Breaking barriers: Integrating Padlet to introduce environmental education to pre-service teachers. *PANYONARA: Journal of English Education,* *7*(1), 77–102. https://doi.org/10.19105/panyonara.v7i1.17849

Roslan, R., Nishio, Y., & Jawawi, R. (2022). Analyzing English language teacher candidates' assessment literacy: A case of Bruneian and Japanese universities. *System*, *111*, 102934. https://doi.org/10.1016/j.system.2022.102934

Saputra, E. R., Hamied, F. A., & Suherdi, D. (2020). The development of beliefs and practices of language assessment literacy: Does a professional learning community help? *Journal of Education for Teaching,* *46*(3), 414–416. https://doi.org/10.1080/02607476.2020.1761250

Shafiee Rad, H., Alipour, J., Jafarpour, A., & Hashemian, M. (2024). Unlocking the magic of digital adaptivity: Unleashing students' writing skills and self-determination motivation through digital literacy instruction. *System*, *125*, 103454. https://doi.org/10.1016/j.system.2024.103454

Shahid, M. K., Zia, T., Bangfan, L., Iqbal, Z., & Ahmad, F. (2024). Exploring the relationship of psychological factors and adoption readiness in determining university teachers' attitude on AI-based assessment systems. *The International Journal of Management Education*, *22*(2), 100967. https://doi.org/10.1016/j.ijme.2024.100967

Swiecki, Z., Khosravi, H., Chen, G., Martinez-Maldonado, R., Lodge, J. M., Milligan, S., … Gašević, D. (2022). Assessment in the age of artificial intelligence. *Computers and Education: Artificial Intelligence*, *3*, 100075. https://doi.org/10.1016/j.caeai.2022.100075

Taghizadeh, M., & and Ejtehadi, A. (2023). Investigating pre-service EFL teachers' and teacher educators' experience and attitudes towards online interaction tools. *Computer Assisted Language Learning*, *36*(8), 1633–1667. https://doi.org/10.1080/09588221.2021.2011322

Viberg, O., Mutimukwe, C., Hrastinski, S., Cerratto-Pargman, T., & Lilliesköld, J. (2024). Exploring teachers' (future) digital assessment practices in higher education: Instrument and model development. *British Journal of Educational Technology*, *55*(6), 2597–2616. https://doi.org/10.1111/bjet.13462

Wulandari, M., & Purnamaningwulan, R. A. (2024). Exploring Indonesian EFL pre-service teachers' experiences in AI-assisted teaching practicum: Benefits and drawbacks. *LLT Journal: A Journal on Language and Language Teaching, 27*(2), 878–894. https://doi.org/10.24071/llt.v27i2.8690

Xi, X. (2023). Advancing Language Assessment with AI and ML–Leaning into AI is Inevitable, but Can Theory Keep Up? *Language Assessment Quarterly*, *20*(4–5), 357–376. https://doi.org/10.1080/15434303.2023.2291488

Xie, H., Chu, H.-C., Hwang, G.-J., & Wang, C.-C. (2019). Trends and development in technology-enhanced adaptive/personalized learning: A systematic review of journal publications from 2007 to 2017. *Computers & Education*, *140*, 103599. https://doi.org/10.1016/j.compedu.2019.103599

Yalcin Arslan, F. (2019). Reflection in pre-service teacher education: Exploring the nature of four EFL pre-service teachers' reflections. *Reflective Practice*, *20*(1), 111–124. https://doi.org/10.1080/14623943.2018.1564652

Zaim, M., Arsyad, S., Waluyo, B., Ardi, H., Al Hafizh, Muhd., Zakiyah, M., … Hardiah, M. (2024). AI-powered EFL pedagogy: Integrating generative AI into university teaching preparation through UTAUT and activity theory. *Computers and Education: Artificial Intelligence*, *7*, 100335. https://doi.org/10.1016/j.caeai.2024.100335

Zhang, L., Zheng, Y., & Yang, D. (2023). Exploring Chinese university English teachers' language assessment literacy: A mixed-method study. *Asia Pacific Journal of Education*, *0*(0), 1–19. https://doi.org/10.1080/02188791.2024.2354684