BRU ELT JOURNAL
Vol. 3 No. 3 (September-December) 2025
ISSN: 2822-1311 (Online)
DOI: https://doi.org/10.14456/bej.2025.18

Integrating Artificial Intelligence Tools in Classroom Language Assessment: Uses, Perceptions, and Pedagogical Implications

Niño G. Costramos[1]
[1]Master of Arts in English Education, Bicol University Open University
Legazpi City, Albay, Philippines
E-mail: ninogarcia.costramos@bicol-u.edu.ph

Received: August 23, 2025
Revised: November 13, 2025
Accepted: November 21, 2025

Abstract

This study explores the uses, perceptions, and pedagogical implications of integrating Artificial Intelligence (AI) tools in classroom language assessment among 30 Junior High School English Language teachers from public and private schools. Using a descriptive-quantitative design, data were collected through a structured survey capturing teachers' experiences and insights. Results indicate that automated scoring software is the most commonly used AI tool, valued for its efficiency, consistency, and ability to provide prompt feedback, while less frequently used tools, such as AI speech labs, reveal opportunities to enhance oral language assessment. Teachers identified advantages including engagement, reduced bias, and adaptive learning, alongside challenges such as overreliance on technology, limited human interaction, and reduced capacity to capture nuanced student performance. Findings suggest that AI integration enhances assessment efficiency, supports higher-order thinking through authentic and adaptive tasks, and fosters multifaceted evaluation, while also reshaping teachers' roles and highlighting the need for professional competence and ethical awareness. Effective AI implementation requires balancing technological tools with human judgment to ensure fairness, meaningful learning, and holistic assessment.

Keywords: artificial intelligence, language assessment, classroom technology, teacher perceptions, pedagogical implications

Introduction

Designing language assessments for language learning remains a persistent challenge for many teachers. The range of assessment formats and tools often creates uncertainty about which approach best fits a specific learning context. These assessments include traditional formats such as multiple-choice, matching, true- false, cloze, dictation, essay tests, and oral interviews, as well as computer-based adaptive tests and online performance tasks. Recent technological advances have introduced tools that record speech samples, analyze written output with natural language processing, and deliver automated scoring at scale (Song, Lee & Jiao, 2025).

Languages assessment is the process of determining an individual's degree of language ability by use of standardized tests and observations (Winna & Sabarun, 2023). The shift from norm-referenced models to performance-based and formative assessment has prompted educators to concentrate on what learners can do with the language rather than on comparative ranks. This shift has created new questions about alignment between assessment and curriculum, the validity of automated measures for complex communicative tasks, and the role of teachers in interpreting assessment data (Bulut et al., 2024).

Generative and other AI technologies have recently accelerated interest in automated scoring and feedback, adaptive testing, and AI-assisted formative assessment. Research has shown promising reliability for some automated scoring systems when compared to human raters, while other studies report limits in contextual judgment and potential bias in automated decisions (Cui & Liang, 2024; Uyar & Büyükahıska, 2025).

Existing literature on AI in language assessment has tended to focus on large-scale testing systems and algorithmic scoring, which leaves classroom-level practice underexamined. Systematic reviews and recent studies call for more empirical research on how teachers implement AI tools in everyday classroom assessment, how teachers perceive the fairness and transparency of these tools, and what professional development teachers require to use AI responsibly. This gap is particularly salient in Asia, where classroom conditions, assessment policies, and edtech adoption vary widely and where few empirical studies have documented teachers' classroom practices and perceptions in national or regional contexts (Boonchom, Piyanukool, & Prachanant, 2024).

The regional and local context gives urgency to this study. Several recent investigations of AI in Asian ELT settings have highlighted teacher concerns about workload, academic integrity, and role re-definition when AI tools become part of classroom practice. Specific studies that focus on public and private schools indicate that English language teachers express both cautious optimism about AI's potential and anxiety about assessment validity and professional accountability (Sakmiankaew et al., 2024; Khaengkhan et al., 2025). These findings suggest that research framed at the classroom level can provide actionable insights for teacher training, curriculum alignment, and policy guidance in Thailand and similar Asian contexts.

This study therefore addresses a concrete gap in the field. There is little empirical evidence exists about which AI tools teachers employ in classroom language assessment, how teachers perceive the advantages and disadvantages of those tools, and how AI integration affects assessment design and pedagogical practice in ELT settings. To address this problem, the study identified the AI tools most commonly used in classroom language assessment. It also examined teachers' perceptions of the benefits and limitations of these tools. Finally, the study analyzed the pedagogical and evaluative implications of AI integration in classroom assessment.

The findings of this study are expected to support classroom assessment strategies, to guide professional development that addresses teacher needs for interpretation and ethical use of AI-generated data, and to provide relevant evidence that policymakers and teacher educators can use to shape guidelines on the pedagogical use of AI in language assessment (Boonchom et al., 2024).

BRU ELT JOURNAL
Vol. 3 No. 3 (September-December) 2025
ISSN: 2822-1311 (Online)

Research Objectives

This study explores the uses, perceptions, and pedagogical implications of integrating Artificial Intelligence (AI) tools in classroom language assessment.

Specifically, this study seeks to:

1. Identify the Artificial Intelligence (AI) tools used in classroom-based language assessment.
2. Examine the advantages and disadvantages of using AI tools in language assessment practices.
3. Analyze the pedagogical implications of AI integration in classroom-based language assessment and language teaching.

Review of Related Literature

Language assessment is a central component of English language teaching, serving as both a measure of learner proficiency and a guide for instructional decision-making. Effective assessment goes beyond testing discrete linguistic elements such as grammar or vocabulary; it evaluates how well learners can use language meaningfully across different contexts (Weigle, 2015). Recent frameworks emphasize the shift from *assessment of learning* to *assessment for learning*, highlighting the formative potential of assessment in supporting learner growth (Fulcher, 2015). Chapelle and Voss (2021) note that such approaches integrate both summative and formative purposes, allowing teachers to use assessment results to adapt instruction and improve student outcomes.

Within this evolving landscape, the emergence of Artificial Intelligence (AI) has introduced significant changes in how language assessments are designed, administered, and interpreted. AI can be defined as the simulation of human intelligence by machines capable of performing complex decision-making and problem-solving tasks (Russell & Norvig, 2021). In education, AI-driven tools are increasingly used to automate scoring, analyze student responses, and provide individualized feedback. In language assessment, these technologies are used to process large volumes of linguistic data and evaluate learner performance with greater precision and efficiency (He & Yu, 2023).

The development of AI in assessment can be traced back to early automated scoring systems such as ETS's *e-rater*, which demonstrated that machine learning models could analyze coherence, grammar, and vocabulary in student writing (Attali & Burstein, 2006). Over time, newer applications such as *Write & Improve* by Cambridge and *Grammarly* have expanded these capabilities for formative classroom use, enabling students to receive immediate and detailed feedback on their writing (Ranalli, Link & Chukharev-Hudilainen, 2022). Similar innovations have appeared in oral assessment, with tools like *Duolingo English Test* and *Pearson's Versant* employing speech recognition and natural language processing to evaluate pronunciation, fluency, and accuracy (Chapelle & Chung, 2015). These AI-based systems are now being integrated into classroom assessment practices, offering teachers and students more dynamic and personalized feedback mechanisms.

Several studies highlight the pedagogical benefits of AI-enhanced assessment. According to Kukulska-Hulme et al. (2022), AI enables real-time, individualized feedback and enhances the consistency and reliability of scoring. Teachers can also use AI-generated data to diagnose specific learner needs and adapt their instruction accordingly, thereby promoting more differentiated teaching. Zhang and Lin (2024) add that AI feedback systems can reduce teacher workload by automating repetitive grading tasks, allowing educators to focus on qualitative aspects of learning such as creativity and critical thinking. These advantages suggest that AI can complement, rather than replace, traditional assessment practices, creating a more balanced and data-informed learning environment.

Nonetheless, the increasing reliance on AI tools in assessment raises important concerns. Scholars have questioned the transparency and fairness of automated scoring systems, as well as the potential for algorithmic bias embedded within training data (Chapelle, 2021; Williamson & Piattoeva, 2022). He and Yu (2023) argue that while AI can enhance efficiency and objectivity, it must be implemented alongside human judgment to ensure valid interpretations of learner performance. Data privacy, accountability, and ethical considerations remain central to ongoing debates about the responsible use of AI in education. These issues underscore the need for teachers to critically understand both the capabilities and limitations of AI-driven assessment tools.

The integration of AI into classroom-based assessment has also made formative assessment more interactive and adaptive. Adaptive testing platforms, for instance, modify question difficulty in real time based on a learner's responses, providing a more accurate measure of proficiency (Lu & Lim, 2023). Natural language processing applications now assist teachers in analyzing students' written and spoken language, producing feedback aligned with communicative learning goals (Ranalli et al., 2022). Through these innovations, assessment becomes more continuous, personalized, and supportive of learner development, aligning closely with communicative and learner-centered pedagogies (Kukulska-Hulme et al., 2022).

Despite these global advances, empirical studies exploring AI-assisted assessment in ELT contexts remain limited. Most existing research focuses on large-scale testing systems rather than classroom-based assessment practices. For instance, while AI tools are increasingly used in high-stakes international exams, relatively few studies have investigated how EFL teachers in Southeast Asian countries adopt and adapt these tools to suit local educational settings (Phumeechanya & Wannapiroon, 2020). The local challenges of infrastructure, teacher readiness, and ethical guidelines also shape how effectively AI can be integrated into classroom practice.

The review collectively illustrates the transformative potential of AI in enhancing the validity, reliability, and efficiency of language assessments. However, gaps remain in understanding how teachers use AI tools for classroom-based evaluation, how they perceive their pedagogical value, and what implications these technologies have for language teaching and learning. This study addresses these gaps by examining the use of AI tools in classroom language assessment, exploring teachers' perceptions of their advantages and limitations, and analyzing the pedagogical and evaluative implications of integrating AI into ELT practice.

BRU ELT JOURNAL
Vol. 3 No. 3 (September-December) 2025
ISSN: 2822-1311 (Online)

Methodology

This study employed a descriptive–qualitative design to provide a detailed understanding of the use of Artificial Intelligence (AI) tools in classroom-based language assessments. The approach focused on identifying the AI tools used in language assessment, examining teachers' perceived advantages and disadvantages, and exploring how AI integration influences language teaching and assessment practices.

The participants were 30 English language teachers from public and private Junior High Schools in Masbate Province, Philippines. Purposive sampling was used to select teachers with relevant experience and familiarity in using AI applications for language instruction and assessment, ensuring that their perspectives would provide meaningful insight into the pedagogical implications of AI integration.

Data were collected through an online structured survey administered between March and April 2025. The instrument comprised Likert-scale items to assess teachers' perceptions of advantages and disadvantages of AI tools, multiple-response questions to identify the most commonly used AI tools, and closed-ended items to capture perceived pedagogical implications.

The collected data were analyzed using a combination of thematic and descriptive quantitative methods. Qualitative responses were examined thematically to identify patterns, insights, and emerging trends regarding AI use in language assessment. Quantitative data from Likert-scale and multiple-response items were analyzed using frequency counts and percentages, which complemented the thematic analysis by highlighting prominent practices and perceptions. This integrated approach provided a comprehensive understanding of how AI tools influence assessment practices and pedagogical decisions in English language teaching across Junior High Schools in Masbate.

Results

The data gathered by the researcher was interpreted, analyzed, and discussed in this study.

Using AI Tools in Classroom Language Assessment

The study investigated the types of AI tools currently employed by 30 Junior High School English Language teachers in public and private schools in Masbate Province. The survey allowed respondents to select multiple tools, reflecting the reality that teachers often combine several technologies to assess different language skills. This approach captures a more accurate picture of classroom practices, as many educators do not rely on a single tool but use multiple AI applications to support varied assessment needs.

Table 1 shows the AI tools used in classroom language assessments. The enumerated AI tools were obtained from the survey conducted by the researcher.

BRU ELT JOURNAL
Vol. 3 No. 3 (September-December) 2025
ISSN: 2822-1311 (Online)

Table 1
Use of AI Tools in Classroom Language Assessments

Rank	Artificial Intelligence Tools	Frequency	Percentage
1	Automated Multiple-Choice Scoring Software	24	80%
2	Automated Essay Scoring Software	22	73.3%
3	Automated Assessment Generators	21	70%
4	AI-Powered Games	18	60%
5	Adaptive Test Tools	7	23.3%
6	AI Speech Labs	2	6.7%

Note: *Percentages exceed 100% cumulatively because respondents could select more than one AI tool. Each percentage represents the proportion of teachers who reported using a specific AI tool.*

The data show that Automated Multiple-Choice Scoring Software is the most frequently used AI tool, with 24 out of 30 teachers (80%) reporting its use. This high percentage highlights the popularity of tools that support rapid and consistent grading of multiple-choice items, which are commonly employed in classroom language assessments. Following closely, Automated Essay Scoring Software was reported by 22 teachers (73.3%), indicating that a majority of respondents utilize AI to evaluate written responses efficiently and objectively. Similarly, Automated Assessment Generators were used by 21 teachers (70%), reflecting a trend where educators apply AI tools to design exercises and quizzes tailored to individual learning levels, supporting more customized assessment practices.

AI-Powered Games were reported by 18 teachers (60%), suggesting moderate adoption of gamified approaches for assessment purposes. These tools allow teachers to assess students in a more interactive and engaging manner. In comparison, Adaptive Test Tools and AI Speech Labs were the least frequently reported, with 7 teachers (23.3%) and 2 teachers (6.7%), respectively. These lower percentages indicate that technologies requiring adaptive algorithms or speech recognition are currently used by a minority of teachers, which may reflect limitations in access, technical capacity, or familiarity with such tools.

Advantages and Disadvantages of AI Tools in Classroom Language Assessments

The study explored respondents' perceptions of the advantages and disadvantages using a Likert-scale questionnaire, allowing participants to express varying degrees of agreement. This approach captures subtle differences in how users view advantages, such as learning and efficiency, versus concerns, including technology reliance and implementation challenges. The results highlight overall trends in approval and caution, providing insight into both the benefits and potential limitations of the system.

Table 2 shows the advantages of using AI tools in conducting classroom language assessments based on the survey conducted with English language teachers.

BRU ELT JOURNAL
Vol. 3 No. 3 (September-December) 2025
ISSN: 2822-1311 (Online)

Table 2
Advantages of AI Tools in Classroom Language Assessment

Advantages	Weighted Mean	Adjectival Rating
Speed and Efficiency	4.83	Strongly Agree
Automated Scoring	4.90	Strongly Agree
Reduced Bias	4.77	Strongly Agree
Engaging Learning	5.00	Strongly Agree
Adaptive Learning	4.67	Strongly Agree
Reliability and consistency	4.50	Strongly Agree
Adaptability and personalized learning	4.60	Strongly Agree
Standardized Scoring	4.87	Strongly Agree

A closer look at the data reveals notable patterns. Engaging Learning achieved a perfect score (5.00), highlighting that respondents found the system highly interactive and motivating. This suggests that the system effectively captures learners' attention and supports active participation. Similarly, Automated Scoring (4.90) and Standardized Scoring (4.87) were among the highest-rated aspects, emphasizing the system's perceived fairness, objectivity, and efficiency in assessment.

Speed and Efficiency (4.83) and Reduced Bias (4.77) also scored strongly, showing that users appreciate both the time-saving nature of the system and its ability to minimize subjective evaluation errors. Meanwhile, aspects related to personalization- Adaptive Learning (4.67) and Adaptability and Personalized Learning (4.60)- received slightly lower but still strongly positive ratings, suggesting that while respondents recognize the system's flexibility, there may be room to further enhance its individualized learning features.

Reliability and Consistency (4.50) received the lowest mean score among the advantages, though it still falls within the "Strongly Agree" category. This indicates that respondents view the system as generally dependable, but consistent performance may be an area for continued monitoring and improvement.

The data reflect a clear trend of strong approval across all evaluated aspects, with particular emphasis on engagement and objective assessment features. The slightly lower but still positive ratings for adaptability and reliability suggest potential areas for enhancement, ensuring the system remains both dynamic and dependable. Overall, respondents perceive the system as an effective, fair, and engaging tool that supports learning while maintaining assessment accuracy.

On the other hand, disadvantages of AI tools in language assessment were identified. Table 3 presents the disadvantages of using AI in classroom language assessments as recorded by the researcher based on the experiences of the English language teachers.

BRU ELT JOURNAL
Vol. 3 No. 3 (September-December) 2025
ISSN: 2822-1311 (Online)

Table 3
Disadvantages of Using AI Tools in Classroom Language Assessments

Disadvantages	Weighted Mean	Adjectival Rating
Potential for cheating	5.00	Strongly Agree
Overreliance on Technology	4.80	Strongly Agree
Raises data privacy concerns	3.73	Agree
Lack of authentic human interaction	4.77	Strongly Agree
Lack of Teacher Training	4.83	Strongly Agree
Cost of development and implementation	4.07	Agree
Overlooked nuances	3.63	Agree

The results indicate that respondents generally recognize several challenges associated with the system. Potential for Cheating (5.00) emerged as the most pressing concern, highlighting worries about academic integrity in digital learning contexts. Other highly rated concerns include Lack of Teacher Training (4.83), Overreliance on Technology (4.80), and Limited Human Interaction (4.77), suggesting that respondents value the role of teacher guidance and personal engagement, and are cautious about relying too heavily on technology.

Lower-rated disadvantages- Data Privacy Concerns (3.73), Cost of Development and Implementation (4.07), and Overlooked Nuances (3.63)- though still acknowledged, were considered less critical. This indicates that respondents perceive operational and technical issues as secondary to factors directly affecting learning quality and experience.

In general, the data show a clear trend: concerns are strongest when the drawbacks could negatively impact learning outcomes or the student experience, particularly around cheating, teacher support, and meaningful interaction. Meanwhile, technical limitations and financial considerations are recognized but seen as less urgent. These insights suggest that while the system is broadly appreciated, addressing these key human and pedagogical challenges is essential for successful implementation.

Pedagogical Implications of AI Integration in Classroom-Based Language Assessment and Language Teaching

The following implications are drawn from the respondents' evaluations, reflecting patterns in the reported advantages, disadvantages, and commonly used AI tools. They highlight how these perceptions may influence teaching approaches and classroom learning experiences.

Emphasis on Authentic Tasks and HOTS. The results show that AI tools are highly valued for enhancing engagement, efficiency, and scoring accuracy. These advantages suggest that AI can support the design of assessments that go beyond rote recall, promoting Higher-Order Thinking Skills (HOTS) such as analysis, evaluation, and creativity. However, concerns about potential for cheating indicate that while AI facilitates efficiency, careful task design is necessary to maintain assessment integrity.

Holistic Learning through Meaningful Activities. High ratings for engaging and adaptive learning demonstrate that AI tools can analyze diverse aspects of student performance. This reflects a shift toward holistic evaluation, where assessments are embedded in meaningful learning experiences rather than isolated tests. Disadvantages

like overlooked nuances and limited human interaction highlight that AI captures measurable outputs but may miss subtler dimensions of student learning, underlining the importance of complementary human oversight.

Efficiency and Consistency in Assessment. The high agreement for automated and standardized scoring shows that AI tools are relied upon for efficient and consistent evaluation, especially in language assessments where scoring can be time-consuming and subjective. This aligns with SOP 1 findings, where tools like automated essay graders and scoring platforms were most commonly used. At the same time, disadvantages such as overreliance on technology and lack of teacher training suggest that efficiency gains are maximized only when teachers understand how to interpret and integrate AI outputs.

Multifaceted Approach to Assessment. The diversity of commonly used AI tools such as covering writing, reading comprehension, and grammar practice, illustrates a multifaceted approach to language assessment. Advantages such as adaptability and personalized learning show that AI supports varied learning needs, while limitations like cost of development and overlooked nuances indicate that balancing breadth and depth of assessment remains a challenge.

Supporting Underrepresented Skills. Although AI tools for spoken language are less widely used, they were still recognized as valuable, reflecting an opportunity to address skills that are traditionally harder to assess, such as oral proficiency and pronunciation. The combination of advantages like engaging learning and adaptive feedback with disadvantages like limited human interaction suggests that AI can complement but not fully replace the richness of teacher-led oral assessment.

Shaping New Learning Experiences. The integration of AI introduces new possibilities for classroom learning, offering interactive and adaptive experiences that were less feasible with traditional assessments. The data on engaging learning and adaptive learning underscore how AI can transform teaching and learning processes, though attention to challenges like overreliance on technology and potential cheating remains critical.

Ethical and Responsible Use. Concerns about cheating and data privacy highlight the ethical dimensions of AI integration. These disadvantages, juxtaposed with high appreciation for efficiency and fairness, suggest that ethical considerations are inherent to AI-supported assessment, influencing how teachers design, implement, and monitor AI-enhanced learning activities.

Discussion

The study investigated the types of AI tools used by 30 Junior High School English Language teachers in public and private schools in Masbate Province, revealing that teachers often combine multiple technologies to assess varied language skills. The survey showed that Automated Multiple-Choice Scoring Software, Automated Essay Scoring, and Automated Assessment Generators were the most commonly used tools, reflecting a clear preference for AI applications that streamline grading, standardize evaluation, and support tailored assessment practices. These patterns align with findings that AI enhances efficiency, objectivity, and scalability in classroom assessments (Zhang, 2025; González-Calatayud et al., 2021). Moderate use of gamified tools and low adoption of adaptive test tools or AI speech labs suggests that technologies requiring complex algorithms or specialized skills remain less accessible,

possibly due to limitations in infrastructure, technical support, or teacher familiarity (Nguyen et al., 2024).

The evaluation of advantages revealed a strong consensus among teachers regarding AI's contributions to classroom language assessment. Perfect scores for engaging learning, along with high ratings for automated and standardized scoring, suggest that AI tools are perceived as highly motivating, interactive, and fair. These findings correspond with studies highlighting that AI can enhance engagement, minimize bias, and support adaptive learning strategies (Sahmaniasl, 2025; Adil, 2024). Slightly lower but still positive scores for adaptability and personalized learning indicate that while AI can address individual learner needs, further integration of differentiated instruction strategies may strengthen assessment effectiveness (Fatima, 2025).

Regarding disadvantages, concerns about potential cheating, lack of teacher training, overreliance on technology, and limited human interaction were highly rated. This suggests that while AI provides efficiency and consistency, teachers remain cautious about ethical, pedagogical, and relational dimensions of assessment. Lower-rated concerns such as data privacy, cost, and overlooked nuances, although acknowledged, appear secondary to issues that directly affect learning quality and experience. These patterns echo previous studies emphasizing that ethical considerations and human oversight are essential for effective AI integration in education (UNESCO, 2024; Zhang, 2025).

The integration of AI tools in classroom language assessment represents a transformative shift in teaching and learning practices, positioning technology as a mediator of more complex pedagogical processes rather than merely a grading mechanism. The data indicate that AI enables more dynamic, interactive, and holistic assessment, allowing educators to capture multiple dimensions of student performance, design authentic and cognitively demanding tasks, and provide timely feedback, while freeing time for instructional planning and learner support (Kaebling & Moore, 2021). At the same time, teachers' concerns regarding overreliance on technology, limited human interaction, and potential misuse underscore the tension between efficiency and meaningful engagement, highlighting the need for careful implementation where AI complements rather than replaces teacher judgment (Alam, 2022). The multifaceted adoption of AI tools from automated scoring to gamified and adaptive applications, demonstrates their potential to address diverse language skills and support personalized learning, yet their effectiveness depends on professional competence, ethical vigilance, and contextual sensitivity (González-Calatayud et al., 2021).

Overall, the findings confirm and extend prior research on AI in education by demonstrating that teachers in Masbate employ multiple AI tools in tandem, maximizing both efficiency and engagement while navigating challenges related to ethics, human interaction, and professional preparation. The patterns observed in this study can be interpreted through a constructivist lens, wherein AI serves as a mediating tool that enhances learning experiences but requires active teacher facilitation to ensure meaningful and equitable outcomes (Özdere, 2023).

Recommendations

Based on the findings, it is recommended that schools provide professional development and training opportunities focused on practical use of AI in classroom assessment. This will help teachers interpret AI-generated results accurately, design meaningful and authentic tasks, and integrate AI in ways that complement their teaching rather than replace it.

Given the diversity of AI tools and their varying adoption levels, it is important for schools to prioritize tools that are accessible, easy to use, and aligned with the specific assessment needs of students. Focusing on tools that support multiple language skills such as writing, reading, and speaking, will enable more comprehensive evaluation while addressing gaps in less frequently assessed areas, such as oral proficiency. Schools may also consider gradually introducing advanced or adaptive AI tools once teachers are confident with foundational applications.

Finally, ethical and responsible use of AI should guide all integration efforts. Clear protocols for monitoring academic integrity, safeguarding student data, and balancing AI assistance with human judgment can ensure that technology enhances learning rather than compromising it. By combining teacher preparedness, careful selection of tools, and attention to ethical considerations, AI can be leveraged effectively to improve assessment practices, promote engagement, and support meaningful learning outcomes in the classroom.

Declaration of Conflicting Interest: The author declares no conflict of interest. **Funding**: The author did not receive financial support for the research, authorship, and publication of this article.

References

Adil, J. G. (2024). AI in education: A systematic literature review of emerging trends, benefits, and challenges. *Seminars in Medical Writing and Education.* https://doi.org/10.56294/mw2025795

Alam, A. (2022). Possibilities and apprehensions in the landscape of artificial intelligence in education. *International Journal of Education and Development Using Information and Communication Technology, 18*(2), 125-138.

Attali, Y., & Burstein, J. (2006). Automated essay scoring with e-rater® V.2. *Journal of Technology, Learning, and Assessment, 4*(3), 1-30.

Boonchom, W., Piyanukool, S., & Prachanant, N. (2024). Trends of using artificial intelligence (AI) technologies in research studies of English language teaching. *BRU ELT Journal.* https://doi.org/10.14456/bej.2024.3

Bulut, O., Beiting-Parrish, M., Casabianca, J. M., Slater, S. C., Jiao, H., & Song, D. (2024). The rise of artificial intelligence in educational measurement: Opportunities and ethical challenges. *arXiv.* https://arxiv.org/abs/2406.18900

Chapelle, C. A. (2021). Challenges in using AI for language assessment. *Language Testing, 38*(3), 361-371. https://doi.org/10.1177/02655322211014357

Chapelle, C. A., & Chung, Y. R. (2015). The promise of NLP and speech processing technologies in language assessment. *Language Testing, 27*(3), 301-315.

Chapelle, C. A., & Voss, E. (2021). *The Routledge handbook of second language acquisition and language testing.* Routledge.

Cui, Y., & Liang, M. (2024). Automated scoring of translations with BERT models: Chinese and English language case study. *Applied Sciences, 14*(5), 1925. https://doi.org/10.3390/app14051925

Fatima, K. (2025). The role of AI in teacher professional development: Implications for AI literacy and training. *AI EDIFY Journal, 2*(1).

Fulcher, G. (2015). *Practical language testing.* Routledge.

González Calatayud, V., Prendes Espinosa, P., & Roig Vila, R. (2021). Artificial intelligence for student assessment: A systematic review. *Applied Sciences, 11*(12), 5467. https://doi.org/10.3390/app11125467

He, L., & Yu, S. (2023). Artificial intelligence in language testing and assessment: Developments, challenges, and future directions. *Language Testing in Asia, 13*(1), 1-15. https://doi.org/10.1186/s40468-023-00209-4

Kaelbling, M., & Moore, C. (2021). Computational intelligence and computing applications. In *2021 International Conference on Computational Intelligence and Computing Applications (ICCICA)* (pp. 1-8). IEEE. https://doi.org/10.1109/ICCICA52458.2021.9697272

Khaengkhan, A., Chaiyarak, S., Koednet, A., Thitipetchakul, C., Skunhom, V., & Iamsen, S. (2025). Language competencies for Thai teachers in the AI era. *ASEAN Journal of Education, 11*(2). https://so01.tci-thaijo.org/index.php/AJE/article/view/274928

Kukulska-Hulme, A., Lee, H., Norris, L., & Zhang, Y. (2022). AI for language learning: Benefits and challenges. *ReCALL, 34*(3), 227-243. https://doi.org/10.1017/S0958344022000039

Lu, X., & Lim, J. M. H. (2023). Adaptive testing and AI in language education: Opportunities and limitations. *System, 116,* 102924. https://doi.org/10.1016/j.system.2023.102924

Nguyen, T. L., Nguyen, H. L., & Le, D. H. (2024). Teachers' perspectives on AI-driven Quillionz for generating EFL reading comprehension quizzes. In *Proceedings of the AsiaCALL International Conference* (Vol. 6, pp. 20-34). https://doi.org/10.54855/paic.2462

Özdere, M. (2023). The integration of artificial intelligence in English education: Opportunities and challenges. *Language Education and Technology, 3*(2).

Phumeechanya, N., & Wannapiroon, P. (2020). A conceptual framework for applying artificial intelligence in adaptive learning for Thai higher education. *International Journal of Emerging Technologies in Learning, 15*(9), 20-36. https://doi.org/10.3991/ijet.v15i09.12619

Ranalli, J., Link, S., & Chukharev-Hudilainen, E. (2022). Automated writing evaluation for formative assessment: Teacher and student perspectives. *Language Learning & Technology, 26*(1), 1-23.

Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.

Sahmaniasl, R. (2025). EFL educators' perceptions of AI-driven assessments: A systematic review. *International Journal of Multidisciplinary Research and Growth Evaluation, 6*(03), 672-680.

Sakmiankaew, I., Nguyen, M. T., Nguyen, N. D. D., & Buripakdi, A. (2024). EFL tertiary teachers' and students' conceptualizations and challenges of using AI tools to improve writing skills in Thailand and Vietnam during the Covid-19 pandemic. *rEFLections, 31*(3), 1120-1143. https://doi.org/10.61508/refl.v31i3.277354

Song, D., Lee, W.-C., & Jiao, H. (2025). Exploring LLM autoscoring reliability in large-scale writing assessments using generalizability theory. *arXiv.* https://arxiv.org/abs/2507.19980

Uyar, A. C., & Büyükahıska, D. (2025). Artificial intelligence as an automated essay scoring tool: A focus on ChatGPT. *International Journal of Assessment Tools in Education, 12*(1), 20-32. https://doi.org/10.21449/ijate.1517994

UNESCO. (2024). *AI competency framework for teachers.* UNESCO.

Weigle, S. C. (2015). *Assessing writing.* Cambridge University Press.

Williamson, B., & Piattoeva, N. (2022). Education governance and datafication: The role of AI in assessment policy. *Learning, Media and Technology, 47*(2), 213-226. https://doi.org/10.1080/17439884.2021.1942957

Winna, W., & Sabarun, S. (2023). The language assessment in teaching-learning English. DIAJAR: Jurnal Pendidikan dan Pembelajaran, 2(4), 413-419. https://doi.org/10.54259/diajar.v2i4.1894

Zhang, X. (2025). Fairness and effectiveness in AI-driven educational assessments: Challenges and mitigation strategies. *Journal of Innovation and Development, 11*(1), 7-10. https://doi.org/10.54097/adcxze34

Zhang, Y., & Lin, Y. (2024). Integrating AI feedback tools in English writing assessment: A teacher perspective. *TESOL Quarterly, 58*(1), 55-78.

Author

Niño G. Costramos holds a Bachelor's degree in Secondary Education specializing in English. He is a Master of Arts in English Education student at Bicol University Open University in Legazpi, Albay. He is from Masbate Province, and once taught at Dr. Emilio B. Espinosa Sr. Memorial State College of Agriculture and Technology- College of Education.