

# Assessment in the Age of Artificial Intelligence: An Autoethnographic Account of Language Teacher in a State University

Jovito B. Madeja

PhD, Eastern Visayas State University, Tacloban City. Email: jovito.madeja@evsu.edu.ph

## Abstract

This study explores language assessment practices in the context of artificial intelligence (AI) through an autoethnographic lens. Conducted at Eastern Visayas State University, the research examines the lived experiences of a language teacher navigating the challenges and opportunities posed by AI-assisted student work. Using reflective journals, teaching artifacts, field notes, narrative vignettes, and document analysis, data were collected over one academic term and analyzed through thematic narrative analysis. Findings reveal the emergence of “detective dilemma,” where the teacher struggles to ensure authenticity and academic integrity amid AI-generated outputs. The study identifies key issues such as the erosion of authorial voice, the “mirage of quality” in AI-assisted writing, and the limitations of AI detection tools. In response, the researcher reconstructs assessment practices by shifting from product-based to process-oriented approaches, including in-class writing, process-based portfolios, and oral assessments. The study further highlights a transformation in teacher identity, from evaluator to facilitator of authentic learning experiences and ethical AI engagement. It concludes that while AI enhances efficiency and accessibility, it cannot replace the human elements of language learning, such as personal voice, cultural context, and critical reflection. The study contributes to the growing discourse on AI in education by providing a contextualized, experience-based perspective on assessment reform in higher education.

**Keywords:** *Artificial Intelligence, Language Assessment, Autoethnography, Academic Integrity, Higher Education.*

## INTRODUCTION

The rapid advancement of artificial intelligence (AI), particularly generative AI technologies, has significantly transformed the landscape of higher education. In recent years, tools such as large language models have become widely accessible to students, reshaping how knowledge is produced, accessed, and assessed. In the field of language education, assessment practices—especially those involving writing, reading, and critical discourse—have been profoundly influenced by these technological shifts. As scholars argue, AI has not only enhanced instructional support but also challenged conventional approaches to evaluating student learning outcomes (Zhao et al., 2024; Mpolomoka, 2025).

Assessment, as a core component of teaching and learning, plays a crucial role in measuring students’ competencies, guiding instruction, and ensuring academic standards. In language education, assessment is particularly complex, as it involves evaluating communicative competence, creativity, and authenticity of expression. Bibliometric analyses of academic writing assessment reveal its centrality in higher education and its evolving nature in response to technological innovations (Nakrowi et al., 2025). However, the emergence of AI tools capable of generating human-like text raises critical questions about the validity and reliability of traditional assessment methods.

The integration of generative AI in higher education has sparked widespread debates on authorship, originality, and academic integrity. Studies indicate that students increasingly use AI tools to support academic tasks, including writing assignments, which complicates the distinction between student-generated and AI-assisted work (Pérez-Pérez et al., 2026). Consequently, educators are confronted with the challenge of designing assessments that accurately reflect students' learning while accommodating or regulating AI use.

Recent systematic and scoping reviews highlight that AI has introduced both opportunities and challenges in educational assessment. On one hand, AI can enhance efficiency, provide immediate feedback, and support personalized learning experiences. On the other hand, it poses risks related to academic dishonesty, over-reliance on technology, and the erosion of critical thinking skills (Weng et al., 2024; Khlaif et al., 2024). These dual implications necessitate a rethinking of assessment frameworks and pedagogical strategies in higher education.

Despite the growing body of literature on AI in education, there remains limited understanding of how teachers themselves experience and navigate these changes in real classroom contexts. Empirical studies have shown that educators are often uncertain about how to adapt their assessment practices in response to AI technologies, leading to varied and sometimes inconsistent approaches (Farazouli et al., 2024). This highlights a gap between theoretical discussions and lived teaching experiences.

Moreover, much of the existing research adopts quantitative or mixed-method approaches, focusing on student outcomes, technological effectiveness, or institutional policies. While these studies provide valuable insights, they often overlook the subjective, reflective, and contextual dimensions of teaching practice. In particular, there is a scarcity of research that captures the personal and professional experiences of language teachers as they engage with AI in assessment.

This gap underscores the need for qualitative approaches that foreground educators' voices and lived experiences. Autoethnography, as a research design, offers a powerful lens to examine the intersection of personal experience and broader cultural and institutional contexts. By situating the researcher within the study, autoethnography enables a nuanced exploration of how assessment practices are shaped, challenged, and transformed in the age of AI.

In the context of a state university, where diverse student populations and resource constraints intersect with emerging technologies, understanding assessment practices becomes even more critical. Language teachers in such institutions often navigate complex realities, balancing institutional expectations, student needs, and ethical considerations in the use of AI. Yet, these contextualized experiences remain underrepresented in current scholarship.

Therefore, this study seeks to address this gap by providing an autoethnographic account of language teaching practices amid the integration of artificial intelligence in assessment. Specifically, it aims to explore how a language teacher experiences, reflects on, and adapts assessment strategies in response to AI use among college students. By doing so, the study contributes to a deeper understanding of the pedagogical, ethical, and practical implications of AI in language education.

Ultimately, this research aspires to inform more responsive, reflective, and context-sensitive assessment practices in higher education. By bridging personal narrative and scholarly inquiry, it offers insights that may guide educators, policymakers, and researchers in reimagining assessment in the age of artificial intelligence.

## Research Questions

- 1) What challenges do I encounter in ensuring authenticity and academic integrity in AI-assisted student outputs?
- 2) How do I adapt my assessment strategies in response to students' use of artificial intelligence in language learning tasks?

## LITERATURE REVIEW

The emergence of artificial intelligence (AI), particularly generative AI, has significantly reshaped higher education, influencing teaching, learning, and assessment practices. Recent literature highlights the rapid expansion of AI-related research, with thousands of Scopus-indexed publications emerging between 2022 and 2025, reflecting the growing importance of AI in educational contexts (Ng & Ho, 2025). This surge underscores how AI is no longer a peripheral tool but a transformative force affecting core academic processes, including assessment.

In higher education, assessment has traditionally been grounded in human judgment, emphasizing authenticity, critical thinking, and demonstration of learning outcomes. However, generative AI tools such as large language models have disrupted these conventions by enabling students to produce high-quality outputs with minimal effort. A scoping review by Xia et al. (2024) revealed that AI is transforming assessment at multiple levels, student, teacher, and institutional necessitating a shift toward more holistic and integrity-focused approaches. This transformation challenges educators to reconsider not only what is assessed but also how learning is evidenced.

In the context of language education, AI introduces both opportunities and complexities. Language assessment, which traditionally evaluates communicative competence, discourse construction, and linguistic creativity, is particularly vulnerable to AI-generated outputs. Jang and Sawaki (2025) argue that AI expands the scope of language assessment by enabling the evaluation of complex linguistic features, such as discourse interaction and speech patterns, while also raising concerns about validity, fairness, and interpretability. These developments suggest that AI is redefining the constructs underlying language proficiency assessment.

At the same time, AI offers pedagogical advantages that can enhance assessment practices. Studies indicate that AI-driven systems can provide immediate feedback, support personalized learning, and improve students' higher-order thinking skills, including creativity and critical analysis (Suazo-Galdames & Chaple-Gil, 2025). Such capabilities align with learning-oriented assessment approaches, which emphasize continuous feedback and student engagement in the learning process.

Despite these benefits, the integration of AI into assessment raises significant ethical and practical concerns. Issues related to academic integrity, authorship, and trust have become central in contemporary discourse. Lelescu et al. (2025) found that educators' adoption of AI is often hindered by concerns about reliability and trust, highlighting the tension between technological innovation and pedagogical accountability. Similarly, concerns about plagiarism and over-reliance on AI challenge the credibility of traditional assessment methods.

Empirical studies also reveal that teachers are actively experimenting with new assessment strategies in response to AI. These include the use of oral examinations, reflective writing, and process-based assessments to ensure authenticity. However, there is no consensus

on best practices, and educators often face uncertainty in adapting to AI-driven changes (Jauhiainen et al., 2025). This lack of standardization reflects the evolving and complex nature of assessment in the AI era.

Furthermore, research focusing on teachers' perspectives indicates varying levels of readiness and acceptance of AI in assessment. Khlaif et al. (2024) found that faculty adoption of AI tools is influenced by factors such as perceived usefulness, ease of use, and institutional support. While some educators embrace AI as a pedagogical tool, others remain skeptical due to concerns about its implications for learning and assessment integrity.

Notably, while existing studies provide valuable insights into AI integration, most adopt quantitative, experimental, or survey-based approaches. Systematic reviews of AI in education emphasize technological effectiveness, policy implications, and student outcomes, often overlooking the nuanced, lived experiences of teachers navigating these changes (Castillo-Martínez et al., 2024; Mustafa et al., 2024). As a result, there is limited understanding of how educators personally interpret, negotiate, and adapt their assessment practices in real-world contexts.

This reveals a significant gap in the literature: the absence of deeply reflective, experience-based accounts of language teachers engaging with AI in assessment. While studies have explored faculty perceptions and institutional responses, few have employed qualitative approaches such as autoethnography to capture the dynamic, context-specific realities of teaching in the AI era. In particular, there is a lack of research situated in state universities, where contextual factors such as resource constraints and diverse student populations may shape assessment practices differently.

In response to this gap, the present study adopts an autoethnographic approach to examine the researcher's lived experiences as a language teacher navigating AI-integrated assessment. By foregrounding personal narrative and reflective inquiry, this study contributes to the literature by providing a contextualized, insider perspective on how assessment practices are transformed in the age of AI. It extends existing research by bridging the gap between theoretical discourse and actual classroom practice, offering insights that are both pedagogically relevant and culturally grounded.

Ultimately, this study contributes to the ongoing reimagining of assessment in higher education by highlighting the complexities, tensions, and possibilities associated with AI integration. It underscores the need for more reflective, adaptive, and human-centered approaches to assessment that acknowledge both the affordances and limitations of artificial intelligence.

## METHODOLOGY

### Research Design

This study employed an autoethnographic research design, a qualitative approach that integrates autobiographical narrative with cultural and institutional analysis. Autoethnography enables the researcher to critically examine personal teaching experiences while situating these within broader educational contexts. In this study, the researcher assumes the dual role of participant and analyst, reflecting on lived experiences in assessing language learning amid the integration of artificial intelligence (AI).

Autoethnography is particularly appropriate for this inquiry as it allows for an in-depth exploration of evolving assessment practices shaped by emerging technologies. It foregrounds subjective experiences, reflexivity, and meaning-making, thereby providing rich, contextualized insights into how AI influences pedagogical decisions, assessment strategies, and teacher identity. Through systematic self-reflection and narrative inquiry, the study captures the complexities and tensions encountered in real classroom settings.

### **Locale of the Study**

The study was conducted at Eastern Visayas State University (EVSU), a public higher education institution in the Philippines. EVSU serves a diverse student population and offers a range of undergraduate and graduate programs, including teacher education and language-related courses.

The university provides a relevant context for this study, as it reflects the realities of state universities where access to digital technologies, including AI tools, is rapidly increasing. Within this setting, language teachers encounter varying levels of student engagement with AI, making it an appropriate site for examining how assessment practices are adapted in response to technological changes.

### **Participants**

The primary participant of this study is the researcher, a language teacher at EVSU, who documents and analyzes personal experiences in designing and implementing assessment practices in the age of AI. As an autoethnographic study, the researcher's lived experiences serve as the central data source.

To enrich the contextual understanding, selected student interactions and outputs may be included as supplementary data. However, students are not treated as direct participants but as part of the educational environment that shapes the researcher's experiences. Any references to students are anonymized to ensure confidentiality.

### **Research Instrument**

Central to this process was the use of an autoethnographic research instrument in the form of a structured reflective journal guide, which served as the primary tool for eliciting, organizing, and deepening the researcher's narratives. The journal guide contained open-ended prompts designed to capture critical dimensions of the researcher's lived experiences, including: (1) experiences in designing and implementing assessment tasks in the presence of artificial intelligence (AI); (2) observations of students' use of AI in language learning activities; (3) challenges encountered in ensuring authenticity and academic integrity; (4) strategies and adjustments made in response to AI integration; and (5) reflections on the evolving role and identity of the teacher. These prompts ensured consistency and depth in the reflective entries while allowing flexibility for spontaneous insights and emergent themes.

Using this instrument, the researcher maintained regular reflective journals documenting personal experiences, insights, and challenges in assessing student work in the presence of AI tools. These entries captured immediate reactions, critical incidents, and evolving perspectives. In addition, teaching artifacts such as assessment tasks, rubrics, instructional guides, and feedback samples were collected to illustrate changes in assessment design and implementation. Field notes were also recorded during and after classroom interactions, focusing on students' use of AI, their engagement with assessment tasks, and classroom dynamics. Significant teaching episodes or "critical moments" related to AI use in assessment

were reconstructed into narrative vignettes to provide rich and detailed accounts of lived experiences. Furthermore, selected student outputs, particularly those suspected or acknowledged to involve AI assistance, were examined to contextualize assessment challenges and pedagogical responses. The data collection process was iterative and recursive, allowing the researcher to continuously reflect on and refine interpretations throughout the study.

### **Data Gathering Procedure**

Data were collected over the course of one academic term through multiple reflective and documentary sources. The researcher maintained regular reflective journals to document personal experiences, insights, and challenges encountered in assessing student work in the presence of artificial intelligence (AI) tools. These journal entries captured immediate reactions, critical incidents, and evolving perspectives. In addition, teaching artifacts such as assessment tasks, rubrics, instructional guides, and feedback samples were gathered to illustrate shifts in assessment design and implementation.

Field notes were also recorded during and after classroom interactions, focusing on students' use of AI, their engagement with assessment tasks, and the overall classroom dynamics. Significant teaching episodes or "critical moments" related to AI use in assessment were reconstructed into narrative vignettes to provide rich, detailed, and evocative accounts of lived experiences. Furthermore, selected student outputs, particularly those suspected or acknowledged to involve AI assistance, were examined to contextualize emerging assessment challenges and corresponding pedagogical responses. The data collection process was iterative and recursive, enabling continuous reflection and refinement of interpretations throughout the duration of the study.

### **Analysis of Data**

Data were analyzed using thematic narrative analysis, an approach that integrates storytelling with systematic coding and interpretation. The process began with data familiarization, wherein the researcher carefully reviewed all collected materials, including journals, artifacts, and field notes, to identify recurring experiences and patterns. This was followed by inductive coding to uncover significant themes related to assessment practices, AI integration, and teaching experiences. The identified codes were then organized into broader thematic categories such as authenticity in assessment, academic integrity, adaptation of assessment strategies, teacher agency, and technological mediation. Through a process of narrative construction or restorying, these themes were synthesized into coherent narratives that reflect the evolution of the researcher's assessment practices and lived realities in the AI era. Finally, the narratives were interpreted in relation to existing theories and literature to provide deeper insights into how personal experiences intersect with broader educational and cultural contexts. This analytical approach ensured that the study remained both rigorous and reflective, effectively balancing personal narrative with scholarly interpretation.

### **Ethical Considerations**

Ethical principles were carefully observed throughout the study to ensure integrity and respect for all individuals involved. Although the primary participant is the researcher, informed consent was obtained from students whenever their outputs or classroom interactions were referenced. Confidentiality and anonymity were strictly maintained through the use of pseudonyms and the removal of identifying information from all data sources. The researcher also practiced reflexivity and transparency by acknowledging potential biases and critically examining personal assumptions and interpretations. All collected data were securely stored

and used exclusively for academic purposes. Moreover, the principle of non-maleficence was upheld, ensuring that no harm, discomfort, or disadvantage was inflicted upon students or any stakeholders as a result of the research process.

## RESULTS AND DISCUSSION

### The “Detective” Dilemma: Challenges in Authenticity and Integrity

The first theme that emerged from the autoethnographic analysis is the persistent tension between maintaining academic integrity and confronting the invisible yet pervasive presence of artificial intelligence (AI) in student outputs. This tension reflects what may be described as a “detective dilemma,” wherein the teacher assumes the role of investigator, attempting to discern authorship in increasingly ambiguous texts. Within the context of Eastern Visayas State University, this challenge is amplified by the growing accessibility of generative AI tools among students.

A central issue identified in this study is the erosion of the authorial voice, a key construct in language education that reflects individuality, cultural grounding, and critical engagement. Reflective journal entries revealed recurring instances of essays that were structurally sophisticated yet lacked cultural authenticity, particularly in topics related to Philippine literature. These findings resonate with recent research indicating that AI-generated texts often exhibit linguistic fluency but lack contextual depth and experiential grounding (Xia et al., 2024). Such outputs create what may be described as a “synthetic coherence” texts that appear academically sound but are detached from the learner’s lived experience.

This phenomenon contributes to what the researcher termed “pedagogical paranoia,” a state of uncertainty and mistrust in evaluating student work. The difficulty in distinguishing between genuine and AI-assisted outputs aligns with the findings of Lelescu et al. (2025), who emphasized that trust remains a major barrier to the adoption of AI in education. When educators cannot confidently attribute authorship, the integrity of assessment practices is fundamentally compromised.

Closely related to this is the theme of the “mirage of quality,” wherein AI-assisted outputs present a polished surface that obscures limited comprehension. Students may submit essays demonstrating advanced vocabulary and syntactic complexity without a corresponding depth of understanding. This supports the argument of Suazo-Galdames and Chaple-Gil (2025) that while AI can enhance the form of student outputs, it does not necessarily guarantee cognitive engagement or higher-order thinking. As a result, traditional assessment criteria particularly those focused on grammar and organization—become insufficient indicators of learning.

The study also highlights the limitations of AI detection tools, leading to what is termed the “detection trap.” Initial reliance on such tools revealed issues of inaccuracy, particularly false positives for non-native English speakers. This finding is consistent with broader critiques that AI detection technologies lack reliability and may disproportionately disadvantage certain student populations (Farazouli et al., 2024). The ethical implications are significant, as accusing students based on probabilistic outputs undermines the principles of fairness and due process in education.

A reflective journal entry encapsulates this dilemma: “I found myself grading the machine rather than the student. The feedback I gave felt like I was shouting into a void, unsure if the recipient was a human mind or an algorithm.” This highlights a critical shift in the nature of assessment from evaluating human learning to interrogating technological mediation.

### **From Product to Process: Reconstructing Assessment Strategies**

In response to these challenges, the findings reveal a significant transformation in assessment practices, characterized by a shift from product-oriented to process-oriented evaluation. This aligns with the concept of learning-oriented assessment, which emphasizes continuous engagement, feedback, and evidence of learning processes rather than solely final outputs (Suazo-Galdames & Chaple-Gil, 2025).

One major adaptation involved the reintroduction of in-class writing tasks, particularly for high-stakes assessments. By conducting writing activities within the classroom, the researcher was able to directly observe students' language production processes, thereby ensuring authenticity. This approach echoes the findings of Farazouli et al. (2024), who noted a resurgence of controlled, in-person assessments as a response to AI-assisted academic work. The "in-class reversion" represents a pragmatic strategy to mitigate AI misuse while preserving the validity of assessment.

Another key innovation was the implementation of process-based portfolios, wherein a significant portion of the grade (60%) was allocated to stages of writing, including outlines, drafts, and peer-review documentation. This approach shifts the focus from the final product to the trajectory of learning, making it more difficult for AI to substitute the entire process. Importantly, students who used AI tools were required to submit their interaction logs and provide reflective accounts of how AI contributed to their work. This practice repositions AI from a tool of academic dishonesty to a pedagogical resource, consistent with emerging perspectives that advocate for the ethical integration of AI in education (Jauhiainen et al., 2025).

The introduction of viva voce or oral defense further strengthened the assessment framework by incorporating real-time interaction. Through brief oral check-ins, students were asked to explain specific elements of their written work, such as vocabulary choices or argumentative structures. This method proved effective in revealing the depth of student understanding and authorship. As noted by Jang and Sawaki (2025), AI-driven changes in education are prompting a return to more interactive and performance-based forms of assessment that capture communicative competence more authentically.

Collectively, these pedagogical pivots demonstrate a move toward multi-modal and human-centered assessment, where authenticity is established through triangulation of evidence rather than reliance on a single output.

### **Discussion: Reimagining the Teacher's Identity**

The findings of this autoethnographic study suggest that the integration of AI in education necessitates a fundamental reimagining of the teacher's role. Rather than functioning primarily as an evaluator of student outputs, the teacher emerges as a curator of learning experiences and a facilitator of ethical engagement with technology.

The challenges identified in Research Question 1 underscore that academic integrity in the AI era is no longer confined to detecting misconduct but involves cultivating a culture of transparency, accountability, and critical awareness. This perspective aligns with recent scholarship emphasizing the importance of AI literacy as a core competency in higher education (Ng & Ho, 2025). In this context, the teacher's role extends beyond assessment to guiding students in the responsible and reflective use of AI tools.

The adaptations described in Research Question 2 highlight the effectiveness of human-centric assessment approaches, which prioritize interaction, reflection, and process over static outputs. By emphasizing oral communication, iterative drafting, and reflective practice, these strategies circumvent the limitations of AI-generated text and re-center the human dimension of learning.

Moreover, the study reveals a contextual insight specific to Eastern Visayas State University: while AI has the potential to support students with limited linguistic proficiency, it also introduces the risk of a new digital divide. Students who use AI as a learning aid may develop their skills, whereas those who rely on it as a shortcut may experience superficial learning. This duality reflects broader concerns in the literature regarding equity and access in AI-enhanced education (Castillo-Martínez et al., 2024).

Ultimately, this study underscores the need for a balanced and reflective approach to AI integration in language assessment. While AI can generate grammatically correct and stylistically polished texts, it cannot replicate the authentic voice, lived experience, and critical consciousness that define meaningful language use. As such, the role of the teacher evolves into fostering these uniquely human dimensions of learning while guiding students to engage with AI as a tool for enhancement rather than substitution.

## CONCLUSION

This autoethnographic study explored the evolving landscape of language assessment in the age of artificial intelligence (AI), situated within Eastern Visayas State University. The findings reveal that AI has fundamentally disrupted traditional assessment practices by blurring the boundaries of authorship, authenticity, and academic integrity. The emergence of AI-assisted student outputs has created what is described as the “detective dilemma,” where the teacher is challenged to distinguish between genuine student work and machine-generated text. This has led to a state of pedagogical uncertainty, reinforcing the need to rethink conventional assessment paradigms.

The study further concludes that AI-generated outputs often create a “mirage of quality,” presenting linguistically polished texts that may lack depth of understanding and personal engagement. This underscores the inadequacy of product-based assessment in capturing authentic learning. Additionally, the limitations and ethical concerns associated with AI detection tools highlight the need for more reliable and equitable approaches to ensuring academic integrity.

In response to these challenges, the study demonstrates that shifting from product-oriented to process-oriented assessment is a viable and effective strategy. Approaches such as in-class writing, process-based portfolios, and oral assessments provide more authentic evidence of student learning and reduce reliance on AI-generated outputs. These strategies align with human-centered assessment practices that emphasize interaction, reflection, and the development of critical thinking skills.

Ultimately, the study reveals that the role of the language teacher is being redefined. Rather than functioning solely as an evaluator of student outputs, the teacher emerges as a facilitator of learning experiences, a guide in ethical AI use, and a promoter of authentic student voice. AI, while powerful, cannot replace the human dimensions of learning—such as personal experience, cultural context, and critical reflection. Thus, the integration of AI in education

calls for a balanced approach that leverages its benefits while preserving the integrity and humanity of language learning.

## RECOMMENDATIONS

Based on the findings and conclusions of the study, the following recommendations are proposed:

### 1. For Language Teachers

Language teachers should adopt process-oriented and human-centered assessment strategies, such as portfolios, oral defenses, and in-class writing activities. These approaches help ensure authenticity and provide a more accurate representation of student learning. Teachers should also integrate AI literacy instruction, guiding students on the ethical and responsible use of AI as a learning tool rather than a shortcut.

### 2. For Higher Education Institutions

Institutions, including Eastern Visayas State University, should develop clear policies and guidelines on AI use in assessment. These policies should address issues of academic integrity, transparency, and ethical use while promoting innovation in teaching and learning practices.

### 3. For Curriculum Developers

Curriculum designers should incorporate AI-responsive assessment frameworks that emphasize critical thinking, creativity, and reflective learning. Tasks should be designed in ways that require personalization, contextualization, and demonstration of understanding beyond what AI can easily generate.

### 4. For Professional Development Programs

Universities should provide continuous faculty training on AI integration, focusing on assessment redesign, ethical considerations, and digital pedagogy. This will empower educators to confidently navigate the challenges posed by AI in the classroom.

### 5. For Future Researchers

Future studies should explore multi-participant perspectives, including students and other educators, to provide a more comprehensive understanding of AI's impact on assessment. Comparative studies across institutions and disciplines are also recommended to validate and expand the findings of this research.

## References

- 1) Castillo-Martínez, I. M., Flores-Bueno, D., Gómez-Puente, S. M., & Vite-León, V. O. (2024). AI in higher education: A systematic literature review. *Frontiers in Education*, 9, 1391485. <https://doi.org/10.3389/feduc.2024.1391485>
- 2) Farazouli, A., Cerratto-Pargman, T., Bolander-Laksov, K., & McGrath, C. (2024). Hello GPT! Goodbye home examination? An exploratory study of AI chatbots' impact on university teachers' assessment practices. *Assessment & Evaluation in Higher Education*. <https://doi.org/10.1080/02602938.2024.2334456>

- 3) Jang, E. E., & Sawaki, Y. (2025). Advancing language assessment for teaching and learning in the era of the artificial intelligence (AI) revolution: Promises and challenges. *Language Testing*, 42(4). <https://doi.org/10.1177/02655322251348685>
- 4) Jauhainen, J. S., Gagagorry Guerra, A., Nylén, T., & Mäki, S. (2025). Generative AI in assessing written responses of geography exams: Challenges and potential. *Journal of Geography in Higher Education*. <https://doi.org/10.1080/03098265.2025.2593484>
- 5) Khlaif, Z. N., et al. (2024). AI-assisted assessment in higher education: A systematic review. *Education Sciences*, 14(10), 1090. <https://doi.org/10.3390/educsci14101090>
- 6) Lelescu, A., Sava, S., Grosseck, G., & Mălita, L. (2025). Exploring trust in generative AI for higher education institutions: A systematic literature review focused on educators. *Humanities and Social Sciences Communications*, 12, 1961. <https://doi.org/10.1057/s41599-025-06253-1>
- 7) Mpolomoka, D. L. (2025). Utilizing artificial intelligence for assessment in higher education. *Pedagogical Research*.
- 8) Mustafa, M. Y., Tlili, A., Lampropoulos, G., et al. (2024). A systematic review of literature reviews on artificial intelligence in education (AIED): A roadmap to a future research agenda. *Smart Learning Environments*, 11, 59. <https://doi.org/10.1186/s40561-024-00350-5>
- 9) Nakrowi, Z. S., Winata, N. T., & Lumettu, A. (2025). Two decades of academic writing assessment in higher education: A bibliometric and technological trend analysis of Scopus (2000–2025). *International Journal of Learning, Teaching and Educational Research*.
- 10) Ng, S.-L., & Ho, C.-C. (2025). Generative AI in education: Mapping the research landscape through bibliometric analysis. *Information*, 16(8), 657. <https://doi.org/10.3390/info16080657>
- 11) Pérez-Pérez, I., González-Afonso, M. C., Plasencia-Carballo, Z., & Pérez-Jorge, D. (2026). Transparency mechanisms for generative AI use in higher education assessment: A systematic scoping review (2022–2026). *Computers*, 15(2), 111.
- 12) Suazo-Galdames, I. C., & Chaple-Gil, A. M. (2025). AI-driven assessment systems in higher education: Effectiveness for enhancing critical thinking and creativity. *Ingénierie des Systèmes d'Information*, 30(6). <https://doi.org/10.18280/isi.300624>
- 13) Weng, X., Xia, Q., Gu, M., & Rajaram, K. (2024). Assessment and learning outcomes for generative AI in higher education: A scoping review on current research status and trends. *Australasian Journal of Educational Technology*.
- 14) Xia, Q., Weng, X., Ouyang, F., Lin, T. J., & Chiu, T. K. F. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. *International Journal of Educational Technology in Higher Education*, 21, 40. <https://doi.org/10.1186/s41239-024-00468-z>
- 15) Zhao, J., Chapman, E., & Sabet, P. G. P. (2024). Generative AI and educational assessments: A systematic review. *Educational Research Perspectives*.