Available online:
https://journal.unismuh.ac.id/index.php/exposure
Exposure: Jurnal Pendidikan Bahasa Inggris


# A NEEDS ANALYSIS FOR DESIGNING A MODULE-BASED AUTOMATED WRITING EVALUATION (AWE) TOOL IN INDONESIAN EFL ACADEMIC WRITING

**Marhamah**[1], **Hendri Yawan**[2], **Litha Nesidekawati Dakka**[3]
[1,2,3] Universitas Sembilanbelas November Kolaka, Indonesia
*Marhamahudri@gmail.com*
Received: September 26, 2025 Revised: April 22, 2026 Accepted: May 18, 2026

**ABSTRACT**
*The growing use of AI Automated Writing Evaluation (AWE) tools in higher education has raised concerns about students' reliance, lack of critical engagement, and ethical uncertainties. However, limited research has examined how AI-supported writing is experienced in Indonesian EFL higher education contexts, particularly regarding its pedagogical and ethical implications in academic writing classrooms. This study, as part of a broader Design-Based Research (DBR) project, focused on a needs analysis through a qualitative case study to explore students' and lecture's perceptions of AI-supported academic writing in an Indonesian EFL context and to inform the design of a digital writing module. Data were collected through interviews with one instructor, focus group discussions with 15 students, and the analysis of student writing drafts. Reflexive thematic analysis was applied to capture patterned meanings and interpretive insights. Findings showed that while AI tools improved grammatical accuracy and text coherence, they often encouraged limited argumentation, homogenized structures, and confusion about academic integrity. These results suggest the importance of embedding scaffolding for reflection, feedback literacy, and digital ethics into instructional design.*
**Keywords:** *Academic Writing, Automated Writing Evaluation, AWE, Ethics*

## INTRODUCTION

The rapid growth of *Artificial Intelligence Automated Writing Evaluation (AI AWE) tools* such as Grammarly, ChatGPT, and Quillbot has brought new possibilities and new challenges to higher education. For many English as a Foreign Language (EFL) students, these tools offer a lifeline that they provide immediate corrections, generate suggestions, and ease the burden of producing accurate academic texts (Gozali et al., 2024). Students who once struggled with grammar or cohesion now have access to feedback that is instant and accessible. Yet, the very strengths of these tools also present problems. Scholars have warned that the convenience of AI can encourage dependency, discourage critical thinking, and blur the lines between assistance and authorship. As Stahl (2025) points out, ethical concerns around originality and integrity are no longer hypothetical but they are part of everyday classroom realities. Weldiningsih (2024) adds that many EFL learners risk letting AI do the thinking for them, which can diminish their own voice in academic writing.

These debates are particularly relevant in Indonesia, where higher education has been under pressure to modernize in line with digital innovation and literacy policies (Aini et al., 2024). In EFL preservice programs, the presence of AI tools is becoming commonplace, often through integration with Learning Management Systems (LMS) or direct use by students in their writing assignments (Anisah et al., 2024; Saifullah et al., 2024). Students report that Grammarly helps them polish grammar and vocabulary, while ChatGPT provides ideas and draft sections of text. Such practices do reduce the barriers of language learning, but they also show a worrying trend: students tend to accept AI suggestions passively, without reflecting on how these align with academic standards (Nguyen et al., 2024). Teachers are noticing a shift, indicating many describe student work as increasingly uniform, with structures and expressions that resemble AI-generated templates rather than authentic attempts at critical writing (Darwin et al., 2024). These observations echo findings in other EFL contexts across Asia, where AI has changed not just the way students write, but also how teachers evaluate originality.

Despite these developments, research on AI-supported academic writing in Indonesia remains scarce. Much of the existing work focuses on the technical side on what AI can do or how students feel about it, rather than on how it actually reshapes classroom practices or challenges established pedagogies (Helmiatin et al., 2024; Yang & Chen, 2023). Little is known about how students combine AI feedback with their own thinking, or how instructors guide them in navigating ethical grey areas. Dwihadiah et al. (2024) underline a critical issue, indicating the lack of digital ethics frameworks in Indonesian universities leaves both students and teachers unsure about what constitutes acceptable use of AI. This lack of clarity widens the gap between the growing adoption of AI tools and the pedagogical structures needed to use them responsibly.

This study responds to that gap by examining how AI AWE tools are currently used in EFL academic writing and what students and teachers see as their main challenges. The aim is not simply to evaluate the tools, but to understand the learning needs that arise when technology becomes part of the writing process. By analyzing students' experiences and instructors' concerns, the study identifies the kinds of support that could make AI use more reflective, ethical, and pedagogically

meaningful. These insights are then used to inform the design of a digital academic writing module integrated into the LMS, with a focus on balancing technical assistance with critical engagement. In this way, the research contributes to ongoing conversations about how AI can be adopted in higher education without undermining academic integrity (Cotton et al., 2023; Kohnke et al., 2023; Zheng et al., 2023), while at the same time offering a grounded response to the realities of the Indonesian EFL context.

## LITERATURE REVIEW

The increasing presence of Artificial Intelligence (AI) in education has generated both optimism and concern. On the one hand, AI Automated Writing Evaluation (AWE) tools promise to address long-standing difficulties faced by EFL preservice teachers, particularly in producing accurate and fluent academic texts. Tools such as Grammarly provide corrective feedback on grammar and vocabulary, while generative models like ChatGPT offer prompts, explanations, and even full drafts of text (Gozali et al., 2024). Studies consistently report that students perceive these tools as useful for overcoming linguistic barriers and enhancing their confidence in writing (Nguyen et al., 2024; Rafida et al., 2024). In Indonesia, where many students still struggle with academic English, the accessibility of AI support is viewed as a breakthrough that enables participation in academic discourse (Aini et al., 2024).

Yet this optimism is echoed by critiques regarding overreliance and the erosion of critical literacy. Scholars argue that when students accept AI-generated suggestions passively, they miss opportunities to engage with the underlying rules of language or the logic of academic argumentation (Darwin et al., 2024). Research across Asian EFL contexts shows that AI-supported writing often results in surface-level fluency but weak reasoning, with texts appearing polished but lacking depth (Nguyen et al., 2024). Weldiningsih (2024) emphasizes that students may become overly dependent on ChatGPT, leading to writing that is less a product of their own intellectual work and more an assemblage of algorithmic outputs. This aligns with international critiques of AI in higher education, which suggest that while

technology can scaffold language mechanics, it risks undermining the cultivation of independent thought and creativity (Selwyn, 2019; Smith, 2023).

Ethical concerns further complicate the adoption of AI in academic contexts. Dwihadiah et al. (2024) highlight the absence of clear guidelines in Indonesian universities regarding plagiarism, paraphrasing, and acceptable levels of AI use. Without such frameworks, students are left to navigate ethical grey zones on their own, which can lead to unintentional misconduct. Rafida et al. (2024) observed that Indonesian and Taiwanese students frequently use AI for paraphrasing, but often fail to provide appropriate citations, raising questions of originality. These findings echo Stahl's (2025) argument that the ethical issues of AI are not external to the technology but embedded within its very affordances. Teachers share similar concerns, reporting that student work increasingly appears homogenized and suspiciously formulaic, raising doubts about authenticity (Darwin et al., 2024). The situation suggests the need for digital ethics education to accompany technical adoption.

The integration of AI into Learning Management Systems (LMS) is often presented as a solution for embedding these tools within structured pedagogies. Studies have shown that AI-enabled LMS environments can enhance student engagement, collaboration, and efficiency (Anisah et al., 2024; Saifullah et al., 2024). Yang and Chen (2023) also argue that learning analytics within AI-supported LMS can provide insights into student writing patterns, potentially helping teachers identify areas for intervention. However, this promise is limited by persistent challenges in aligning AI functionalities with meaningful pedagogy. Helmiatin et al. (2024) found that while public universities in Indonesia were quick to adopt AI in LMS, many lacked strategies to ensure that these technologies were used for critical learning rather than mechanical correction. Internationally, McNally and Ford (2017) and Zheng et al. (2023) highlight the value of Design-Based Research (DBR) for developing AI-supported curricula, stressing the need for iterative cycles of design, testing, and refinement.

The discussion above indicates a complex interplay between opportunity and risk in the use of AI AWE tools in EFL academic writing. While these tools effectively reduce linguistic barriers and enhance surface-level accuracy, they

simultaneously introduce challenges related to overreliance, ethical ambiguity, and the homogenization of student writing. Existing studies have largely focused on students' perceptions, technological affordances, and general ethical concerns. However, there remains a significant gap in empirical research that examines how AI tools are actually integrated into real classroom practices in the Indonesian EFL context, particularly from both student and instructor perspectives.

More importantly, there is limited research that translates these challenges into pedagogically grounded design implications, especially within a structured instructional framework such as Learning Management Systems (LMS). This gap is critical because without context-sensitive pedagogical interventions, the integration of AI risks becoming purely technical rather than educational. Therefore, this study is urgent as it responds directly to the rapid, yet unregulated, adoption of AI in Indonesian higher education, where clear pedagogical and ethical frameworks are still evolving.

The novelty of this study lies in its integration of needs analysis within a Design-Based Research (DBR) framework, which not only documents current practices but also systematically informs the development of an AI-supported academic writing module. Unlike prior studies that remain descriptive, this research contributes by bridging the gap between empirical classroom realities and instructional design, offering actionable insights for developing reflective, ethical, and critically engaged AI-supported writing pedagogy. In doing so, this study advances the discourse on responsible AI integration by positioning technology not merely as a tool, but as part of a pedagogical ecosystem that must be carefully designed and continuously refined.

## METHODS

This study represents the initial phase of a broader Design-Based Research (DBR) project aimed at developing and evaluating a digital academic writing module integrating AI Automated Writing Evaluation (AWE) tools. The present phase focuses specifically on conducting a needs analysis through a qualitative case study design.

The selection of a qualitative case study approach is theoretically and methodologically grounded. First, the study seeks to explore complex, context-dependent phenomena, namely how students and instructors interact with AI tools in authentic academic writing settings. Such phenomena cannot be adequately captured through quantitative measures alone, as they involve perceptions, beliefs, experiences, and ethical interpretations that are inherently subjective and socially constructed. Recent qualitative research literature emphasizes that interpretive approaches are essential for understanding meaning-making processes in technology-mediated learning environments (Aspers & Corte, 2019; Flick, 2022). A qualitative approach therefore enables an in-depth understanding of these dynamics by capturing rich, descriptive data.

Second, the use of a case study design is particularly appropriate because the research is situated within a specific institutional and pedagogical context. Contemporary case study scholars argue that case studies are especially valuable for investigating bounded systems where context plays a crucial role in shaping behavior and interaction (Yin, 2023; Harrison et al., 2020). In this study, the "case" is defined as the integration of AI tools within an EFL academic writing course at a public university in Indonesia. This bounded focus enables the researcher to uncover nuanced patterns of behavior, such as how students negotiate AI assistance and how instructors interpret emerging challenges.

Third, the qualitative case study aligns closely with the principles of Design-Based Research (DBR). DBR emphasizes iterative cycles of design, implementation, analysis, and refinement, all grounded in real educational contexts (Zheng et al., 2023; Barab & Squire, 2021). Within this framework, needs analysis serves as a critical foundational phase, ensuring that subsequent design decisions are informed by actual user needs rather than theoretical assumptions. By employing a qualitative case study, this research captures the authentic challenges, expectations, and practices of participants, thereby enhancing the ecological validity of the future instructional design.

Moreover, the choice of this design is justified by the exploratory nature of the research problem. Given that AI integration in Indonesian EFL writing classrooms is still emerging and under-researched, there is limited prior knowledge

to guide hypothesis-driven inquiry. Therefore, an inductive, qualitative approach allows patterns and themes to emerge organically from the data, providing a grounded basis for theory development and instructional innovation (Flick, 2022).

To capture diverse perspectives, the study employed multiple sources of qualitative data. Data were gathered through semi-structured interviews with two lectures, focus group discussions with 15 fourth-semester undergraduate EFL students in a public university, and the collection of student writing samples as artefact. Interviews and FGDs were selected because they provide nuanced accounts of participants' lived experiences and allow for probing of emergent themes (Flick, 2022). Student writing drafts, meanwhile, offered supporting evidence of how AI AWE tools were used in practice, enabling triangulation between reported experiences and observable outcomes (Korstjens & Moser, 2019). Participants were recruited purposively to ensure representation of those actively engaged in academic writing courses, with instructors chosen for their responsibility in teaching and assessing student writing. This combination of data sources strengthened the credibility and richness of the needs analysis (Creswell & Poth, 2019).

Data analysis followed the principles of reflexive thematic analysis (Braun & Clarke, 2019), which emphasizes researcher reflexivity and iterative engagement with the data. This approach was particularly suitable as it allows the identification of patterned meanings across diverse data sources while accommodating the interpretive lens of the researcher. Analysis proceeded in several stages: familiarization through repeated reading of transcripts and documents, systematic coding of meaningful segments, generation of initial themes, and refinement of themes through iterative discussion. Reflexivity was maintained throughout, with analytic memos used to capture the researcher's evolving interpretations and potential biases (Korstjens & Moser, 2019). By applying reflexive thematic analysis, the study was able to move beyond surface-level descriptions of AI use to uncover deeper insights into students' dependency, teachers' ethical concerns, and the pedagogical needs that will guide subsequent DBR design cycles.

## RESULTS

The needs analysis revealed three interconnected themes: (1) surface-level reliance on AI for linguistic accuracy, (2) limited critical engagement with AI-generated content, and (3) ethical ambiguities in the use of paraphrasing tools. These findings highlight a tension between the technical improvements AI provides and the intellectual depth that remains underdeveloped in student writing.

First, students reported frequent use of Grammarly to polish grammar and sentence structure, yet they admitted to limited understanding of the corrections offered. As one participant explained, "Saya pakai Grammarly hampir setiap kali menulis. Grammar jadi bagus, tapi saya tidak tahu kenapa salahnya begitu" (M2) (I use Grammarly almost every time I write. The grammar looks fine, but I don't really know why it was wrong in the first place." M2-translation). This illustrates that Grammarly was perceived primarily as an error-fixing mechanism rather than a learning scaffold. Document analysis supported this claim, showing that while students' drafts demonstrated fewer grammatical errors, they often lacked elaboration and critical reflection. The emphasis on linguistic surface features over deeper meaning-making indicates that AI tools may inadvertently reinforce a product-oriented mindset in writing.

Second, students described using ChatGPT to generate ideas or even whole sections of text, often without evaluating its quality or relevance. One student noted,

"ChatGPT membantu saya mencari ide untuk pendahuluan, tapi kadang hasilnya umum sekali dan saya langsung pakai" (M5) — (ChatGPT helps me find ideas for the introduction, but sometimes the results are very general and I just use them as they are) (M5-translation)

Document analysis confirmed this pattern, as introductions across multiple drafts appeared homogenized, with generic phrasing and limited contextual grounding in scholarly literature. This gap between product and process was also evident when students struggled to defend their arguments during classroom discussions. One participant said, "Saat dosen tanya kenapa argumen saya begitu, saya bingung menjawab karena itu sebenarnya hasil AI" (M7) —(When the lecturer asked why my argument was like that, I was confused to answer because it was actually generated by AI) (M-7-translation). These examples highlight how reliance

on AI outputs without reflective engagement weakens argumentation, leaving students unable to articulate or justify their written claims.

Third, ethical concerns emerged most prominently in students' use of Quillbot for paraphrasing. One student shared, "Quillbot sering saya pakai untuk parafrasa, tapi saya tidak yakin itu termasuk plagiarisme atau tidak" — (I often use Quillbot for paraphrasing, but I'm not sure if that counts as plagiarism or not) (M7-translation)". This uncertainty suggests a lack of clarity regarding academic integrity in the context of AI-assisted writing. Document analysis revealed that nearly half of the student drafts contained extended paraphrases without citation, producing texts that appeared original in form but borrowed heavily in substance. Such practices blur the boundary between acceptable academic support and disguised plagiarism, underscoring the need for clearer guidance on ethical AI use.

These student practices were mirrored in the instructor's observations. The lecturer interviewed expressed concern that,

"Bagus memang kelihatan itu tulisannya mahasiswa, tapi argumennya dangkal sekali. Terasa sekali kalau banyak dibantu AI. Ndak terlalu nampak hasil pemikirannya sendiri di situ" (D1) - (The students' writing looks good on the surface, but the arguments are very shallow. It is very obvious that it was heavily assisted by AI. Their own thinking doesn't really show) (D1-translation)

This perception aligns with the document analysis, which consistently showed technically polished yet substantively shallow work. The lecturer also noted the striking similarity across drafts, stating, "Sering sekali saya temukan itu draft mahasiswa, mirip sekali strukturnya. Kayak memang pake template sama semua" (D1) — "I often find student drafts that look very similar in structure, as if they all used the same template." This points to the homogenizing effect of AI, where outputs replicate formulaic structures and obscure students' individual voices. Finally, the lecturer emphasized the urgent need for ethical guidelines, remarking, "Kita perlu panduan yang jelas tentang kapan mahasiswa boleh pakai AI dan kapan harus menulis sendiri" (D1) — "We need clear guidelines about when students are allowed to use AI and when they must write independently." This concern was substantiated by the document analysis, which showed that in the absence of explicit policies, students applied AI inconsistently and often uncritically.

## DISCUSSION

The findings of this study suggest that while AI AWE tools demonstrably enhance the technical quality of student writing, they simultaneously create conditions that undermine reflective learning, critical argumentation, and ethical awareness. This paradox is consistent with earlier studies showing that AI tools often shift attention toward surface-level accuracy rather than deeper cognitive engagement (Weldiningsih, 2024; Nguyen et al., 2024). For example, Grammarly's corrective feedback enabled students to produce texts with fewer grammatical errors, yet the interviews revealed that learners rarely understood why their errors occurred. As one student admitted, they simply accepted corrections without reflection, which supports Darwin et al.'s (2024) observation that AI tools risk reinforcing passive cognitive habits rather than cultivating active critical thinking. This reliance on surface-level correction reflects what Stahl (2025) calls the "affordance dilemma," where technological affordances both enable and constrain meaningful learning.

Furthermore, the findings highlight how ChatGPT was used as an idea generator but often uncritically incorporated into student writing. The homogenized introductions documented in student drafts echo Gozali et al.'s (2024) framework, which cautions that without explicit training in feedback literacy, students tend to accept AI outputs wholesale rather than interrogating them for relevance and accuracy. This lack of evaluative judgment aligns with Yang and Chen's (2023) study of AI-supported academic writing in LMS platforms, which found that students frequently reproduce generic AI-generated content without linking it to disciplinary knowledge or scholarly sources. In the present study, the inability of students to defend their arguments during class discussions demonstrates a troubling disconnect between the appearance of coherent argumentation and the absence of underlying comprehension. This resonates with Darwin et al.'s (2024) concern that AI-enabled writing risks producing texts that are linguistically competent but intellectually hollow.

The ethical ambiguity surrounding Quillbot further complicates the role of AI in academic writing. Students' uncertainty about whether paraphrasing with AI constitutes plagiarism reflects the lack of clear institutional guidelines in Indonesia,

a gap also emphasized by Dwihadiah et al. (2024). The document analysis, which revealed extensive uncited paraphrasing, illustrates what Nguyen et al. (2024) describe as the erosion of academic integrity when students are unsure how to navigate AI-generated text. These concerns are not unique to Indonesia; Rafida et al. (2024) reported similar trends among students in Taiwan, suggesting that paraphrasing tools blur the line between original composition and academic dishonesty. Stahl (2025) argues that such ethical dilemmas are not incidental but intrinsic to the way AI systems are designed to optimize linguistic output rather than intellectual ownership. Thus, addressing ethics in AI-supported writing requires more than prohibitions—it necessitates pedagogical frameworks that explicitly integrate digital ethics into the curriculum (Helmiatin et al., 2024).

The lecturer's concerns about the homogenization of student work shed further light on the pedagogical implications of AI use. The perception that student drafts looked like "templates" produced by AI resonates with broader critiques of how generative AI risks flattening diversity of expression (Selwyn, 2019; Smith, 2023). In writing contexts, this homogenization diminishes students' authorial identity and makes it difficult for instructors to assess originality. Gozali et al. (2024) argue that to counter this effect, feedback literacy and metacognitive engagement must be explicitly scaffolded so that students learn to critically evaluate and adapt AI outputs rather than reproduce them unaltered. The lecturer's call for clear guidelines aligns with Anisah et al. (2024) and Saifullah et al. (2024), who emphasize that AI tools must be embedded within structured LMS frameworks that include transparent policies and guided activities. Without such interventions, as this study shows, students risk perceiving AI as an unquestioned authority, thereby weakening their capacity for independent thought.

These findings affirm the need for a design-based pedagogical response. Design-Based Research (DBR) provides a systematic approach for iteratively designing, testing, and refining instructional interventions grounded in authentic educational challenges (Zheng et al., 2023; Barab & Squire, 2021). In the present context, the identified issues—surface-level reliance on AI, limited critical engagement, ethical uncertainty, and the homogenization of writing—provide an empirical foundation for the development of a digital academic writing module

integrated into the Learning Management System (LMS). More importantly, the findings suggest that AI-supported writing instruction should not merely focus on improving linguistic accuracy, but also on fostering reflective thinking, ethical awareness, metacognitive regulation, and learner autonomy (Gozali et al., 2024; Weldiningsih, 2024).

From a sociocultural learning perspective, AI tools may function as mediational instruments that facilitate collaborative meaning-making, scaffolding, and reflective interaction when integrated within guided pedagogical practices (Sharples, 2023). Nevertheless, in the absence of adequate pedagogical mediation, students may position themselves as passive recipients of AI-generated content rather than active participants in the construction of knowledge. This tendency was evident in the present findings, where students frequently adopted AI-generated ideas and corrections without critically examining their appropriateness or validity. Such patterns potentially weaken students' authorial identity and critical reasoning skills, both of which are fundamental components of academic writing competence (Darwin et al., 2024; Smith, 2023).

The findings may also be interpreted from the perspective of metacognitive learning, which emphasizes learners' capacity to monitor, regulate, and evaluate their own cognitive processes during learning activities (Qin et al., 2022; Sun & Zhang, 2022). Within AI-assisted writing environments, metacognitive engagement becomes particularly crucial because students are required to make informed and reflective decisions regarding the acceptance, modification, or rejection of AI-generated feedback. However, the findings revealed that many students relied heavily on AI outputs without sufficient evaluative judgment, indicating limited development of feedback literacy and self-regulated learning strategies. This observation is consistent with previous studies suggesting that AI-supported writing environments may foster learner dependency when students are not explicitly guided to critically engage with technological feedback (Nguyen et al., 2024; Yang & Chen, 2023).

Therefore, the proposed instructional design should encourage students not only to utilize AI tools effectively, but also to critically assess, adapt, and justify AI-generated feedback and ideas throughout the writing process. Such pedagogical

integration is essential for fostering critical literacy, ethical reasoning, and sustainable learner autonomy within AI-supported academic writing contexts (Dwihadiah et al., 2024; Helmiatin et al., 2024). Accordingly, AI should not be positioned as a substitute for human cognition; rather, it should function as a pedagogical partner that supports the development of reflective, self-regulated, and critically engaged EFL writers (Stahl, 2025; Weldiningsih, 2024).

## CONCLUSION

This study has shown that while AI AWE tools such as Grammarly, ChatGPT, and Quillbot support Indonesian EFL preservice teachers in producing technically accurate academic texts, they also foster surface-level reliance, weaken critical engagement, and generate ethical uncertainties. They tended to accept AI feedback without reflection, incorporated generic content uncritically, and demonstrated confusion about plagiarism when using paraphrasing tools. Document analysis reinforced these concerns, revealing polished but shallow drafts with homogenized structures. Lectures echoed these issues, stressing the urgent need for clear ethical guidelines and pedagogical scaffolding. The findings suggest the necessity of moving beyond technical adoption toward a pedagogically grounded and ethically responsible integration of AI in academic writing.

The study contributes theoretically by extending current discussions on AI-supported writing through the perspectives of sociocultural learning and metacognitive engagement, emphasizing that meaningful AI integration requires guided interaction, reflective learning, and critical evaluation of AI-generated outputs. More importantly, the study offers practical implications for EFL higher education contexts. The findings highlight the importance of integrating AI literacy into EFL curricula to help students develop critical awareness, evaluative judgment, and responsible AI-use practices. In addition, the study underscores the need for clear institutional and instructional guidelines regarding ethical AI use, particularly in relation to paraphrasing, citation, authorship, and academic integrity.

Furthermore, the findings suggest that AI-integrated writing instruction should incorporate reflective and process-oriented learning activities that encourage students to critically assess, adapt, and justify AI-generated feedback rather than

passively accepting it. Such pedagogical practices are essential for fostering learner autonomy, critical literacy, self-regulated learning, and ethical reasoning in AI-supported academic writing environments. Accordingly, AI should be positioned not as a replacement for students' intellectual engagement, but as a pedagogical partner that supports the development of reflective, independent, and critically engaged EFL writers.

However, this study is limited by its focus on a single institutional context and a relatively small group of participants, which restricts the generalizability of the findings. Future research should expand to multiple institutions and employ longitudinal or mixed-methods designs to examine how pedagogical interventions with AI AWE tools influence students' critical engagement and ethical awareness over time.

## ACKNOWLEDGEMENT

The authors want to expand their gratitude to the Ministry of Higher Education, Science, and Technology for funding this research under research contract 071/E5/PG.02.00.PL/2025/116/UN56.D.01/PN.03.00/2025.

## AI DECLARATIONS

We acknowledge that we used ChatGPT-5 during the writing of this manuscript. This tool was employed to help me brainstorm key ideas, refine my language phrasing, and reduce the word count. We significantly adapted and modified the content generated by the AI tool to accurately reflect my unique style and voice and further adjustments were made throughout the writing process.

## REFERENCES

Aini, N., Kurniarahman, I., Widiati, U., Cahyono, B. Y., & Basthomi, Y. (2024). Indonesian university students' perspectives on integrating AIEd into English language learning. *Issues in Educational Research, 34*(3), 803–824. https://www.proquest.com/docview/3104449889

Anisah, L., Yawan, H., & Marhamah, M. (2024). Artificial intelligence enhanced learning management system: Supporting Merdeka Belajar–Kampus Merdeka (MBKM) at a state university in Indonesia. *International Journal of Education and Social Studies Management, 4*(3), 917–931. https://doi.org/10.52121/ijessm.v4i3.414

Aspers, P., & Corte, U. (2019). What is qualitative in qualitative research. *Qualitative Sociology, 42*(2), 139–160. https://doi.org/10.1007/s11133-019-9413-7

Barab, S., & Squire, K. (2021). Design-based research: Putting a stake in the ground. Journal of the Learning Sciences, 30(3), 1–15.

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health, 11*(4), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. Education and Information Technologies, 28(6), 7195–7211. https://doi.org/10.1007/s10639-023-11978-2

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE.

Darwin, D., Rusdin, D., Mukminatien, N., Suryati, N., & Laksmi, E. D. (2024). Critical thinking in the AI era: An exploration of EFL student perception, benefits, and limitations. *Cogent Education, 11*(1), 2290342. https://doi.org/10.1080/2331186X.2023.2290342

Dwihadiah, D. L., Niyu, N., & Purba, H. (2024). Digital ethics model concerning the use of ChatGPT in Indonesian higher education. *Information, Medium, Society: Journal of Public Studies, 23*(1), 1–22. https://doi.org/10.18848/2691-1507/CGP/v23i01/1-22

Flick, U. (2022). The SAGE handbook of qualitative research design. SAGE.

Gozali, I., Wijaya, A. R. T., Lie, A., Cahyono, B. Y., & Suryati, N. (2024). ChatGPT as an automated writing evaluation tool: Feedback literacy development and AWE tools' integration framework. *JALT CALL Journal, 20*(1), 1–20. https://doi.org/10.29140/jaltcall.v20n1.1200

Harrison, H., Birks, M., Franklin, R., & Mills, J. (2020). Case study research: Foundations and methodological orientations. Forum Qualitative Sozialforschung, 21(1).

Helmiatin, H., 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

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. TESOL Quarterly, 57(4), 1105–1118. https://doi.org/10.1002/tesq.3318

Korstjens, I., & Moser, A. (2019). Series: Practical guidance to qualitative research. Part 4: Trustworthiness and publishing. *European Journal of General Practice, 24*(1), 120–124. https://doi.org/10.1080/13814788.2017.1375092

McKenney, S., & Reeves, T. C. (2019). Conducting educational design research (2nd ed.). Routledge.

Nguyen, H. T., Bui, T. T., & Vo, D. T. (2024). The impact of AI writing tools on academic integrity: Unveiling English-majored students' perceptions and practical solutions. *International Journal of TESOL & Education, 4*(1), 110–121. https://doi.org/10.54855/ijte.241110

Qin, C., Zhang, R., & Xiao, Y. (2022). A questionnaire-based validation of metacognitive strategies in writing and their predictive effects on the writing performance of English as foreign language student writers. *Frontiers in Psychology, 13*, 1071907. https://doi.org/10.3389/fpsyg.2022.1071907

Rafida, T., Suwandi, S., & Ananda, R. (2024). EFL students' perception in Indonesia and Taiwan on using artificial intelligence to enhance writing skills. *Jurnal Ilmiah Peuradeun, 12*(3), 987–1016. https://doi.org/10.26811/peuradeun.v12i3.1520

Saifullah, S., Yawan, H., Syafitri, N., & Nurhaliza, S. (2024). Integrating AI chatbot into learning management system: Enhancing student engagement and learning outcomes. *Dharmas Education Journal, 5*(2), 1346–1359. https://doi.org/10.56667/dejournal.v5i2.1605

Selwyn, N. (2019). Should robots replace teachers? AI and the future of education. Polity Press.

Sharples, M. (2023). Towards social generative AI for education: Theory, practices and ethics. arXiv.

Smith, M. (2023). Artificial intelligence and the homogenization of student writing: Risks and responsibilities in higher education. *Journal of Academic Ethics, 21*(2), 211–229. https://doi.org/10.1007/s10805-022-09460-3

Stahl, B. C. (2025). Locating the ethics of ChatGPT—Ethical issues as affordances in AI ecosystems. *Information, 16*(2), 104. https://doi.org/10.3390/info16020104

Sun, Q., & Zhang, L. J. (2022). Understanding learners' metacognitive experiences in learning to write in English as a foreign language: A structural equation modeling approach. *Frontiers in Psychology, 13*, 986301. https://doi.org/10.3389/fpsyg.2022.986301

Weldiningsih, I. (2024). Balancing AI and authenticity: EFL students' experiences with ChatGPT in academic writing. *Cogent Arts & Humanities, 11*(1), 2392388. https://doi.org/10.1080/23311983.2024.2392388

Yang, Y., & Chen, X. (2023). Learning analytics of AI-supported academic writing in LMS: Patterns and challenges. *Frontiers in Education, 8*, 1126083. https://doi.org/10.3389/feduc.2023.1126083

Zheng, L., Bender, S., & Wang, Y. (2023). Design-based research on the development and implementation of an AI course integrating competition and education. *Frontiers of Education in China, 18*(3), 365–389. https://doi.org/10.3868/s110-008-023-0031-5