

Rethinking Assessment in the AI Era: Emerging Challenges and Responses

Laila Noor

The School of Teacher Education, University of Central Florida, USA

Correspondence

Email: laila.noor@ucf.edu

Abstract

The widespread adoption of generative artificial intelligence (AI) has prompted renewed attention to how academic writing should be assessed in higher education. While AI supports multilingual writers throughout the writing process, it also challenges traditional product-based assessment by obscuring students' learning, authorship, and decision-making. Guided by Sociocultural Theory, this qualitative descriptive study explored how assessment can be redesigned for AI-mediated academic writing. Data included classroom reflections, student writing artifacts, multiple drafts, revision records, and AI-supported writing activities. Data was analyzed using Braun and Clarke's (2023) Reflexive Thematic Analysis. Thematic analysis generated six interconnected themes that reconceptualize assessment as the documentation of learning rather than the evaluation of final products alone. The findings suggest that process-oriented assessment, including scaffolded drafting, revision histories, reflective writing, AI-use disclosures, and ongoing feedback, provides richer evidence of students' cognitive engagement, ethical source use, and writing development. The study proposes a process-oriented assessment framework that promotes transparency, responsible AI use, learner accountability, and more authentic assessment in multilingual writing classrooms.

ARTICLE HISTORY

Received: 01 June 2026

Revised: 20 August 2026

Accepted: 01 September 2026

KEYWORDS

Generative Artificial
Intelligence, AI Literacy,
Process-Oriented
Assessment, Academic
Writing

How to cite this article (APA 7th Edition):

Noor, L. (2026). Rethinking assessment in the AI era: Emerging challenges and responses. *Language Testing in Focus: An International Journal*, 14, 1–20. <https://doi.org/10.32038/ltf.2026.14.01>

Introduction

Generative artificial intelligence (GenAI) has rapidly become part of academic writing in higher education. Multilingual writers increasingly use tools such as ChatGPT, Claude, Gemini, Copilot, Grammarly, QuillBot, and Scholarcy to brainstorm ideas, organize arguments, revise drafts, improve language accuracy, paraphrase text, summarize sources, and receive immediate feedback (Selim, 2024; Subedi &

Nyamasvisva, 2024). Rather than functioning simply as writing assistants, these technologies are reshaping how writing is planned, negotiated, and revised throughout the composing process (Barrot, 2024; Dong, 2024; Kostka & McMartin-Miller, 2026). Their growing integration into L2 writing has prompted TESOL scholars to reconsider long-standing assumptions about writing instruction, feedback, authorship, and assessment.

The rapid adoption of AI also presents a fundamental challenge for writing assessment. Product-oriented assessment has traditionally judged the quality of completed texts by emphasizing organization, coherence, language accuracy, and rhetorical effectiveness. Although these criteria remain important, they reveal little about how students generate ideas, evaluate evidence, revise arguments, or respond to feedback while composing. In AI-mediated writing environments, much of this intellectual work occurs before the final draft is submitted, making the completed text an increasingly limited representation of students' learning, cognitive engagement, and authorial decision-making. Consequently, recent scholarship argues that assessment should document learners' reasoning, revision histories, reflective dialogue, and interactions with AI rather than relying exclusively on final written products (Dong, 2026; Pitychoutis & Spathopoulou, 2026; Tian & Louw, 2026).

At the same time, researchers acknowledge that AI-supported writing offers both opportunities and risks. Studies have shown that AI can improve writing quality, language development, vocabulary development, feedback, and revision when learners engage with it critically and receive appropriate instructional guidance (Al-khresheh, 2024; Barrot, 2024; Dong, 2024; Kostka & McMartin-Miller, 2026; Noor & Islam, 2025). Yet these benefits are accompanied by persistent concerns regarding uncertain authorship, hallucinated references, excessive dependence on AI, diminished critical engagement, and evolving conceptions of academic integrity (Kasneci et al., 2023; Mollick, 2024; Pack & Maloney, 2024). Some scholars have further argued that excessive reliance on AI may contribute to cognitive displacement syndrome, where learners increasingly delegate analysis, synthesis, and decision-making to AI rather than developing these capabilities themselves (Schwartz & Diliberti, 2026; Starominski-Uehara, 2026; Zhai, 2022). These concerns have shifted attention beyond whether students use AI toward how they evaluate AI-generated suggestions, verify information, and maintain responsibility for their academic work. Consequently, AI literacy is increasingly understood as involving critical evaluation, source verification, ethical decision-making, and informed human judgment rather than avoidance of AI technologies (Kasneci et al., 2023; Veldhuis et al., 2025).

Despite these advances, important questions remain. Much of the existing literature has focused on learners' perceptions, experimental interventions, writing quality, conceptual discussions, systematic reviews, teacher assessment literacy, or practitioner reflections. Comparatively little qualitative research has examined how advanced English learners engage in AI-mediated academic writing across multiple

drafts, revisions, and reflective activities within authentic classroom settings, or how these processes can inform assessment practices. This gap is particularly important because responsible AI use extends beyond detecting AI-generated text; it involves understanding how learners evaluate sources, make revision decisions, negotiate authorship, and demonstrate accountability throughout the writing process. Addressing this gap, the present qualitative study investigates how writing instructors can reshape assessment practices in AI-mediated academic writing courses to focus on learning processes rather than final products.

The study was guided by the following research question:

RQ1: How can writing instructors reshape assessment practices in AI-mediated academic writing courses to focus on learning processes rather than final products?

Literature Review

The current interest in generative AI builds on more than a decade of research into technology-enhanced language learning rather than emerging as an isolated innovation. Early work focused on Computer-Assisted Language Learning (CALL), automated writing evaluation, natural language processing (NLP), and digital writing assistants that primarily supported grammar, vocabulary, and language accuracy. As AI technologies matured, writing support expanded from rule-based correction to more interactive assistance for planning, revision, and feedback (Barrot, 2024). The emergence of large language models (LLMs), particularly ChatGPT, marked another shift by enabling context-sensitive dialogue and personalized writing support, prompting growing attention to AI literacy, feedback engagement, assessment, and ethical academic writing (Al-khresheh, 2024; Barrot, 2024; Dong, 2024; Pitychoutis & Spathopoulou, 2026).

Current Research Streams in AI-Mediated Academic Writing

Recent scholarship on generative AI in language education has expanded rapidly into several interconnected research streams that collectively shape current understanding of AI-mediated academic writing. One prominent stream examines AI-assisted writing, exploring how large language models support planning, drafting, revision, and language development while raising questions about learner autonomy and higher-order writing skills (Barrot, 2024; Escalante et al., 2023; Mahapatra, 2024; Seo, 2024). A second stream focuses on AI-mediated feedback, investigating how AI-generated feedback influences learner engagement, revision practices, and written corrective feedback. Rather than replacing teacher feedback, these studies increasingly advocate hybrid approaches that combine AI support with human pedagogical guidance (Dong, 2024, 2026; Kostka & McMartin-Miller, 2026; Pitychoutis & Spathopoulou, 2026).

A third line of inquiry addresses AI literacy, arguing that effective AI use requires learners and teachers to develop critical evaluation skills, ethical decision-making, and reflective judgment rather than technical proficiency alone (Barrot, 2024; Dong, 2024;

Tian & Louw, 2026). Parallel to this work, another research stream examines academic integrity, with growing attention to authorship, plagiarism, hallucinated references, transparency, and responsible AI use in academic writing (Al-khresheh, 2024; Kasneci et al., 2023; Mollick, 2024; Pack & Maloney, 2024). More recently, researchers have begun investigating AI-supported assessment, proposing process-oriented approaches that document learners' reasoning, revision histories, and reflective decision-making instead of evaluating only final written products (Tian & Louw, 2026).

Collectively, these research streams suggest that AI-mediated writing is no longer viewed solely as a technological innovation but as a pedagogical, ethical, and assessment issue that requires integrated instructional and research perspectives.

Ethical Source Use in AI-Mediated Academic Writing

Ethical source use has long been recognized as a central component of academic writing, extending beyond the technical application of citation conventions. From an Academic Literacies perspective, source use is understood as a socially situated practice in which writers learn to participate responsibly in disciplinary knowledge construction rather than merely avoiding plagiarism. This perspective shifts attention from citation mechanics to the development of authorial identity, disciplinary voice, and epistemic responsibility.

Generative AI introduces new complexities into these long-standing discussions. Beyond concerns about plagiarism, students must now determine whether AI-generated information is accurate, distinguish authentic scholarship from fabricated or hallucinated references, verify citations against original sources, and decide how AI assistance should be acknowledged. Emerging scholarship increasingly frames these practices as questions of epistemic responsibility, requiring writers to justify knowledge claims and remain accountable for every source incorporated into their work, regardless of whether it originated from AI or traditional research tools (Dong, 2024; Kasneci et al., 2023; Pack & Maloney, 2024). This shift positions ethical source use as a broader issue of critical judgment, transparency, and responsible participation in academic knowledge production, providing the conceptual foundation for the present study.

The Teacher's Role in AI-Mediated Writing and Assessment

Although much of the emerging literature on generative AI has focused on students' writing practices, increasing attention is being given to the role of teachers in shaping responsible AI use. Recent scholarships suggest that effective AI integration depends not only on students' technological competence but also on teachers' ability to mediate learning, design appropriate assessment, and exercise informed professional judgment. Rather than functioning as evaluators of final products alone, instructors are increasingly positioned as facilitators who guide students in critically interpreting

AI-generated content, verifying sources, and making informed writing decisions (Dong, 2026; Kostka & McMartin-Miller, 2026).

This shift also requires expanded forms of teacher expertise. AI literacy enables educators to recognize both the affordances and limitations of generative AI, while assessment literacy supports the redesign of writing assessment to capture drafting, revision, reflection, and responsible AI engagement instead of relying exclusively on polished final texts. Tian and Louw (2026) argue that AI should function as a mediational partner that supports teachers' reflective decision-making rather than replacing professional expertise. Similarly, Kostka et al. (2026) contend that successful multilingual writing instruction depends on continued teacher scaffolding, dialogue, and contextualized feedback, even when AI tools are integrated into the writing process. Emerging evidence therefore points toward a human–AI partnership, where teachers retain responsibility for pedagogical judgment, ethical guidance, and assessment decisions. This perspective reinforces the need to investigate assessment practices that make learning visible while preserving teacher agency in AI-mediated writing classrooms.

Assessment in AI-Mediated Academic Writing

The rapid adoption of generative AI has prompted a fundamental reconsideration of writing assessment. Product-oriented approaches, which primarily evaluate final written texts, provide limited evidence of students' reasoning, revision practices, and engagement with AI throughout the writing process. As AI-generated writing becomes increasingly sophisticated, researchers argue that assessment should capture how learners develop ideas, evaluate AI-generated suggestions, and justify their writing decisions rather than focusing exclusively on polished products (Tian & Louw, 2026). This shift reflects a broader movement toward process-oriented assessment, where drafting, revision of histories, reflective commentaries, and AI-use documentation provide evidence of learning that final submissions alone cannot reveal.

Recent scholarship also suggests that effective assessment depends on strengthening learners' feedback literacy. Students need opportunities not only to receive AI-generated feedback but also to interpret, evaluate, and selectively apply it while maintaining responsibility for their own writing decisions (Dong, 2024, 2026). Reflection therefore becomes an essential assessment practice because it reveals students' metacognitive awareness, critical evaluation of AI-generated content, and evolving understanding of authorship and academic integrity. At the same time, AI-supported assessment continues to rely on human judgment. Rather than replacing teachers' expertise, AI should function as a mediational tool that supports instructors' professional decision-making, contextual interpretation, and formative feedback (Kostka & McMartin-Miller, 2026; Tian & Louw, 2026). Recent work also advocates AI transparency, including revision logs, AI-use disclosures, and documented writing processes, as mechanisms for making learning visible while promoting accountability and ethical AI use. Collectively, these developments position assessment as a process

of documenting learning, reflection, and informed decision-making rather than merely judging the quality of final written products, a perspective that provides the foundation for future research.

Although recent scholarship has advanced understanding of AI-assisted writing, AI literacy, ethical source use, and assessment, these areas have largely been examined independently. Consequently, relatively little qualitative research has explored how writing instructors redesign assessment practices to document multilingual learners' writing processes, revision decisions, and responsible AI use in authentic classroom contexts. As AI becomes embedded in academic writing, understanding how instructors make learning visible while preserving authorship, accountability, and academic integrity has become an important but underexplored area of inquiry. The present study addresses this gap by examining how instructors reshape assessment practices to emphasize learning processes rather than final written products.

Theoretical Framework

This study is grounded in Vygotsky's (1978) Sociocultural Theory (SCT), which views learning as a socially mediated and developmental process shaped through interaction with people, cultural practices, and mediational tools. From this perspective, generative AI should not be regarded simply as a writing assistant but as a mediational resource whose educational value depends on how learners critically engage with it through instructor guidance, peer dialogue, and reflective writing practices. Recent scholarship similarly argues that AI-supported learning is most effective when combined with human judgment, feedback, and opportunities for reflection rather than replacing pedagogical expertise (Dong, 2024, 2026; Kostka & McMartin-Miller, 2026; Tian & Louw, 2026). SCT therefore shifts attention from evaluating final written products to understanding how learning develops through drafting, revision, feedback, and decision-making within the Zone of Proximal Development. This perspective aligns with the present study's process-oriented assessment framework, which conceptualizes writing development as an evolving, socially mediated process in which transparency, reflection, and responsible AI use provide meaningful evidence of learning beyond the finished text.

The conceptual framework proposes that assessment in AI-mediated academic writing should shift from evaluating what students produce to examining how learning develops throughout the writing process. This framework was informed by classroom observations, student writing artifacts, instructor reflections, Sociocultural Theory, and recent scholarship on AI-assisted writing, assessment literacy, feedback literacy, and responsible AI use. Rather than treating generative AI as a threat to assessment, the framework conceptualizes AI as a mediational tool whose educational value depends on transparency, critical evaluation, and informed human judgment.

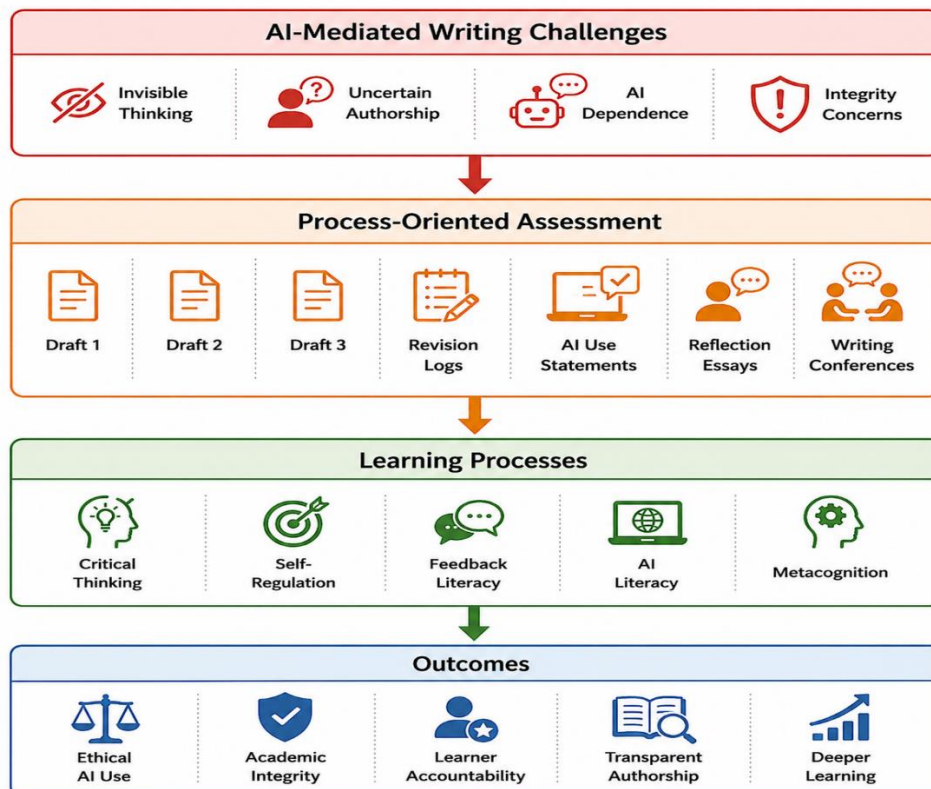
The framework begins with challenges introduced by AI-supported writing, including uncertain authorship, invisible cognitive processes, overreliance on AI-generated

content, limited evidence of revision, and concerns about academic integrity. These challenges suggest that final written products alone no longer provide sufficient evidence of learning. In response, the framework proposes process-oriented assessment practices, including scaffolded drafting, multiple draft submissions, revision histories, AI-use disclosures, reflective commentaries, writing conferences, source verification, and iterative feedback. Together, these practices make students' reasoning, revision decisions, and engagement with AI visible throughout the writing process.

The framework further proposes that documenting these processes strengthens critical thinking, metacognitive awareness, feedback literacy, AI literacy, self-regulation, and ethical source use. In this way, assessment extends beyond judging writing quality to examining how learners evaluate evidence, negotiate AI-generated suggestions, and maintain authorship. The framework therefore positions process-oriented assessment as a practical response to AI-mediated writing, one that preserves human judgment while promoting transparency, responsible AI use, and ethical academic writing.

Figure 1

Conceptual Framework for Process-Oriented Assessment in AI-Mediated Academic Writing



Note. The figure presents the proposed conceptual framework for process-oriented assessment in AI-mediated academic writing. It illustrates how process-oriented assessment practices make students' learning visible by documenting drafting, revision, reflection, and AI-supported decision-making, thereby promoting ethical AI use, transparent authorship, academic integrity, and deeper learning.

Methodology

Research Design

This study employed a qualitative descriptive research design (Sandelowski, 2000, 2010) to investigate how generative AI is reshaping the assessment of academic writing in higher education. Qualitative description was appropriate because the study sought to provide a practice-oriented account of classroom assessment rather than generate new theory. The focus was on documenting instructors' and students' experiences with AI-mediated writing, identifying recurring assessment challenges, and examining instructional responses that emerged in authentic classroom contexts. Using qualitative content analysis, the study explored patterns related to assessment practices, authorship, revision, feedback, and AI use. The analysis generated six interconnected themes that revealed limitations of traditional product-based assessment while pointing toward the value of process-oriented approaches that make students' writing development more visible.

Context and Data Sources

The study was conducted during the 2026 academic year in advanced academic writing courses at a U.S. higher education institution. Participants included 50 advanced English learners and 10 writing instructors. Data was collected from multiple classroom-based sources, including instructors' reflective journals, student writing assignments, successive drafts, revision logs, AI-assisted revision activities, reflective commentaries, instructor feedback records, and course assessment materials. Drawing on multiple sources provided a richer understanding of how students composed, revised, incorporated feedback, and used generative AI throughout the writing process while allowing evidence to be triangulated across different forms of classroom documentation.

Particular attention was given to assignments requiring students to submit an initial draft, one or more revised drafts, and a written reflection explaining how instructor feedback and AI-generated suggestions informed subsequent revisions. These materials enabled examination of both the finished text and the decision-making processes that shaped its development. Rather than evaluating only the final product, the analysis considered how students negotiated authorship, revision, source use, and AI assistance across successive stages of writing, offering a more comprehensive view of writing development and assessment in AI-mediated learning environments.

Data Analysis

Data were analyzed using Reflexive Thematic Analysis (RTA) following Braun and Clarke's six-phase approach (Braun & Clarke, 2006, 2019, 2021a, 2021b, 2023). Consistent with a qualitative descriptive design, the analysis sought to remain close to participants' experiences while developing meaningful interpretations of assessment practices in AI-mediated academic writing. Classroom reflections, writing submissions, revision logs, AI-supported writing activities, and assessment artifacts were read repeatedly to develop familiarity with the dataset. Initial codes were

generated inductively from the data and refined through repeated engagement with individual data extracts and the dataset as a whole. Rather than treating coding as a mechanical process, the analysis was understood as an iterative and interpretive activity in which the researcher actively developed patterns of shared meaning across multiple data sources. Reflexive memos were maintained throughout the analytic process to document methodological decisions, evolving interpretations, and the influence of the researcher's dual role as instructor and researcher. Candidate themes were reviewed, refined, and clearly defined through recursive movement between the coded data, the complete dataset, and the study's conceptual framework. This process resulted in six interconnected themes that informed the development of a process-oriented assessment framework emphasizing scaffolded drafting, revision tracking, reflective writing, AI-use transparency, source verification, and formative writing checkpoints.

Researcher Positionality

The researcher served as both the course instructor and the primary analyst, providing sustained engagement with classroom interactions and students' writing development across the study. This position offered contextual knowledge and insight into how learners negotiated AI-supported writing, revision, and assessment while also shaping the interpretive process. By using Braun & Clarke's Reflexive Thematic Analysis, reflexivity was treated as an ongoing analytic practice rather than an attempt to eliminate researcher influence (Braun & Clarke, 2019, 2021a, 2023). Reflexive memos documented evolving interpretations, methodological decisions, and underlying assumptions throughout the analysis. Although the interpretations were inevitably researcher-mediated, they remained grounded in multiple data sources and informed by the study's qualitative descriptive orientation (Sandelowski, 2000, 2010).

Trustworthiness

Several strategies were employed to establish the trustworthiness of the findings following the criteria of credibility, transferability, dependability, and confirmability proposed by Lincoln and Guba (1985) and operationalized for thematic analysis by Nowell et al. (2017). Credibility was strengthened through data triangulation across classroom reflections, student drafts, revision logs, AI-use documentation, and assessment records. Dependability and confirmability were supported by maintaining a comprehensive audit trail that documented coding decisions, analytic memos, and theme development throughout the analytic process. Researcher reflexivity was maintained through reflective journaling, enabling the researcher to acknowledge underlying assumptions and critically consider how the dual role of instructor-researcher might influence data interpretation. Finally, transferability was supported through thick description, providing detailed contextual information and illustrative examples that enable readers to judge the applicability of the findings to comparable AI-mediated academic writing contexts (Lincoln & Guba, 1985; Nowell et al., 2017).

Findings

Analysis of classroom reflections, student writing samples, revision histories, AI-supported writing activities, and assessment records generated six interconnected themes that describe how generative AI is reshaping both students' writing practices and instructors' approaches to assessment. Rather than emerging as isolated issues, these themes collectively reveal tensions between conventional product-focused evaluation and the increasingly iterative, collaborative nature of AI-mediated writing. Across the data, students' learning, decision-making, source use, and revision processes often remained invisible when assessment emphasized only the final text. The findings therefore suggest that assessing the writing process, rather than the finished product alone, provides a more complete and credible understanding of student learning, authorship, and responsible AI use in multilingual academic writing contexts.

Theme 1: Learning Becomes Invisible When Assessment Privileges Products Over Processes

One of the clearest patterns across the data was that product-based assessment provided only a partial view of student learning. When instructors evaluated only the final essay, they could judge the quality of the completed text but had little evidence of the thinking, decision-making, and revision that shaped it. Classroom reflections suggested that polished writing often concealed the intellectual work behind the final submission. Some students produced well-written essays yet struggled to explain the rationale for their arguments, source selection, or organizational choices during subsequent discussions. This disconnects raised questions about whether the finished product accurately reflected students' own understanding and development as writers.

Variation across drafts reinforced this concern. While some students demonstrated steady refinement of ideas through multiple revisions, others submitted final papers that differed substantially from their earlier drafts. Without access to intermediate versions, instructors found it difficult to determine whether these changes reflected genuine learning, extensive external assistance, or AI-supported text generation. The growing availability of generative AI intensified this challenge because grammatically accurate and well-organized essays no longer served as reliable indicators of students' cognitive engagement.

These findings suggest that assessment focused solely on the final product overlooks essential aspects of the learning process. Making students' thinking visible through draft histories, revision records, and reflective commentaries offers a more complete basis for evaluating writing development and understanding, rather than judging only the finished text.

Theme 2: Authorship Is Negotiated Through Human Judgment Rather Than AI Detection

A second theme concerned the changing nature of authorship in AI-supported writing. Although generative AI assisted students with brainstorming, organization, language refinement, and revision, it also blurred long-standing assumptions about originality and ownership. Instructor reflections revealed that the central question was not whether AI had been used, but how students engaged with it. Many students described AI as a tool for improving organization or strengthening arguments, while others acknowledged relying on AI-generated revisions because the language appeared more sophisticated than their own. From the students' perspective, these practices represented legitimate writing support rather than a transfer of authorship.

Classroom observations, however, exposed important differences in students' engagement with AI-generated content. Some students were able to explain, justify, and adapt AI suggestions to align with their intended meaning, demonstrating continued ownership of their work. Others struggled to discuss ideas presented in their submissions, suggesting that they had accepted AI-generated text with limited reflection or understanding. These cases did not automatically indicate academic misconduct, but they raised legitimate questions about students' cognitive engagement and responsibility for the ideas presented.

Together, the findings suggest that authorship exists along a continuum rather than as a simple distinction between human- and AI-generated writing. Consequently, human judgment informed by students' explanations, revision histories, and reflective documentation offers a more meaningful basis for evaluating authorship than relying on AI detection tools alone.

Theme 3: Revision Makes Learning Visible

A third theme demonstrated that revision transparency made students' learning visible in ways that final papers alone could not. Drafts, revision histories, and revision logs documented how students' thinking developed throughout the writing process, allowing instructors to evaluate learning rather than simply judging finished products. Early drafts often revealed weak organization, limited argument development, citation inaccuracies, and language difficulties. Across subsequent drafts, many students responded to instructor feedback, reflected on their choices, and strengthened both the quality of their writing and the clarity of their ideas.

Revision logs offered particularly rich evidence because they captured students' reasoning behind specific changes. When students explained why they accepted, adapted, or rejected instructor or AI-generated suggestions, instructors gained insight into their critical thinking, metacognitive awareness, and developing authorial judgment. This shifted assessment away from evaluating isolated outcomes toward understanding each student's learning trajectory.

The data also suggested that transparent revision practices supported fairer assessment in AI-mediated writing environments. Documenting how AI was used enabled instructors to evaluate the quality of students' decisions instead of treating AI use itself as evidence of misconduct. Rather than asking whether AI contributed to a paper, instructors could examine how students engaged with AI recommendations and whether those decisions reflected meaningful learning. In general, these process-oriented records reduced uncertainty about authorship while providing a more complete and trustworthy account of student engagement and writing development.

Theme 4: Reflection Documents Cognitive and Ethical Decision-Making

The fourth theme concerned the role of reflection in revealing students' cognitive and ethical decision-making during AI-assisted writing. Reflection essays showed considerable differences in how students understood and managed their writing processes. Some demonstrated strong metacognitive awareness by explaining how instructor feedback informed revisions and why they accepted, adapted, or rejected AI-generated suggestions. Others focused primarily on correcting grammar and vocabulary, suggesting a narrower view of writing as editing rather than knowledge construction.

Instructor reflections indicated that reflective writing often revealed learning that was not evident in the final paper. For instance, students sometimes showed meaningful growth in source evaluation, citation practices, and AI literacy despite only modest improvements in their written products. Requiring students to explain their use of AI also encouraged more deliberate and critical engagement with AI-generated recommendations, reinforcing responsible authorship and academic integrity. Taken together, the findings suggest that reflection serves both instructional and assessment purposes by documenting students' reasoning, self-regulation, and cognitive engagement throughout the writing process rather than evaluating the final product alone.

Theme 5: Transparency Supports Responsible AI Use

The fifth theme demonstrated that process-oriented assessment offers a practical way to promote ethical AI use without relying on surveillance or attempts to eliminate AI from academic writing. Across instructor reflections and assessment records, participants increasingly recognized that evaluating only the final product provided limited insight into students' learning in AI-supported environments. Instead, assessment practices such as draft checkpoints, revision logs, reflective writing, source verification tasks, and AI-use disclosures shifted attention toward the decisions students made throughout the writing process.

This approach encouraged students to document how they interacted with AI, justify revision choices, and take responsibility for the final version of their work. Instructors observed that students were generally more willing to disclose AI use when transparency was treated as part of learning rather than as evidence of misconduct.

These conversations created opportunities to discuss authorship, responsible AI use, and academic integrity in more meaningful ways.

Students also appeared to become more discerning users of AI. Rather than accepting AI-generated suggestions without question, they increasingly verified information, evaluated the relevance of recommendations, and modified outputs to reflect their own ideas and rhetorical intentions. The findings therefore suggest that responsible AI use is better supported through assessment designs that prioritize transparency, reflection, and accountability, allowing instructors to evaluate students' reasoning and decision-making instead of focusing solely on AI detection.

Theme 6: Assessment Shifts from Detecting Misconduct to Documenting Learning

The final theme points to a broader reconsideration of academic integrity in AI-mediated writing. Rather than centering assessment on detecting plagiarism or unauthorized AI use, instructors increasingly questioned whether detection alone could provide a fair or meaningful evaluation of student learning. Classroom reflections suggested that identifying AI-generated text had become both technically unreliable and pedagogically limiting. As a result, instructors placed greater emphasis on documenting how students developed their work instead of attempting to determine whether AI had been used.

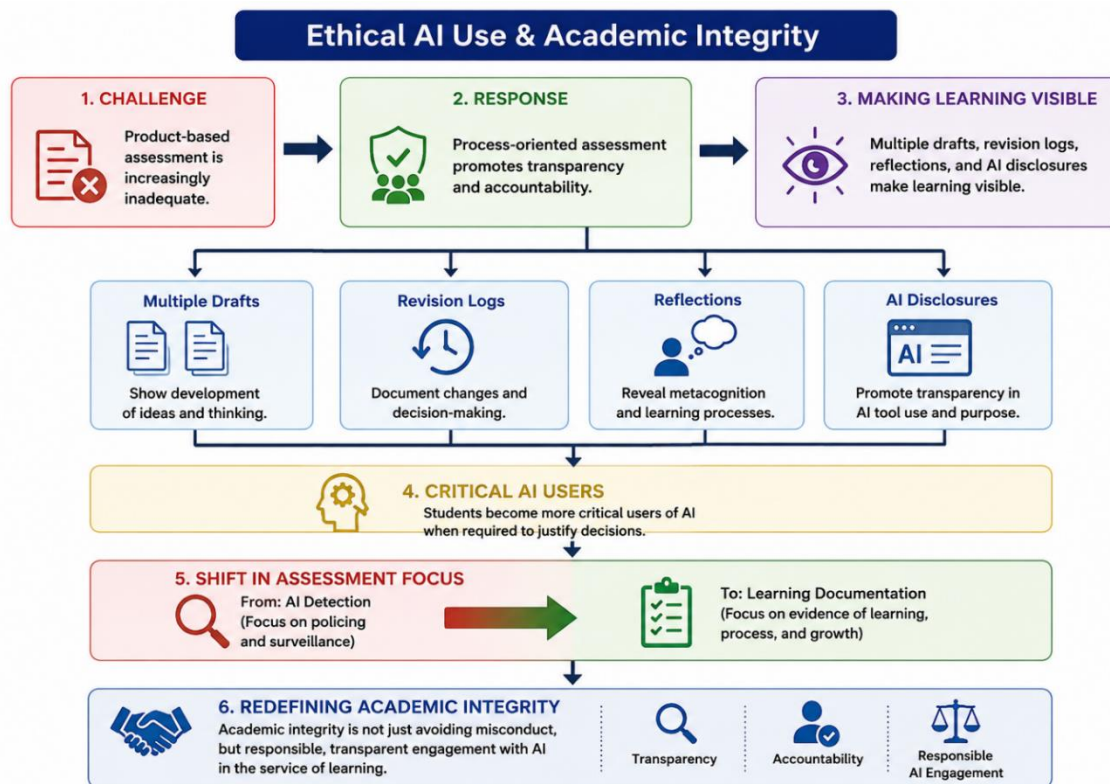
Student submissions indicated that many integrity concerns arose from uncertainty rather than deliberate misconduct. Learners often struggled to distinguish appropriate AI assistance from authorship, evaluate the credibility of AI-generated information, or decide how AI use should be acknowledged. These challenges reinforced the need for explicit instruction and transparent assessment expectations.

Process-oriented assessment addressed these concerns by requiring evidence of students' learning throughout the writing process, including draft submissions, revision of histories, reflective commentaries, and AI-use disclosures. Together, these artifacts provided a richer account of students' reasoning, decision-making, and intellectual growth than the final paper alone.

Across the six themes, the findings suggest that the central challenge of AI-mediated writing is not simply detecting misconduct but designing assessments that capture authentic learning. Assessment practices that value transparency, reflection, iterative revision, and responsible AI engagement offer a more sustainable approach to supporting academic integrity while recognizing how writing evolved in the AI era.

Figure 2

Ethical AI Use and Academic Integrity Through Process-Oriented Assessment



Note. The figure presents the proposed process-oriented assessment framework for AI-mediated academic writing, illustrating how documenting drafting, revision, reflection, and AI use provides richer evidence of learning than evaluating final products alone. Grounded in the study findings and relevant literature, the framework reconceptualizes academic integrity as transparent, accountable, and responsible engagement with AI rather than merely the absence of misconduct.

Discussion

This study contributes to the growing body of research on AI-mediated academic writing by demonstrating that assessment must evolve alongside the changing nature of writing in the era of generative artificial intelligence. Across the six themes, findings consistently emphasized that evaluating final written products alone no longer provides sufficient evidence of students' learning, authorship, critical thinking, or responsible AI use. Instead, the findings suggest that process-oriented assessment offers a more authentic approach by documenting how students generate ideas, evaluate AI-generated suggestions, revise their work, and justify their writing decisions throughout the composing process. Rather than treating AI as a threat to academic integrity, the findings indicate that assessment can be redesigned to make learning visible while encouraging transparency, accountability, and responsible AI engagement.

The findings extend Sociocultural Theory by illustrating that learning in AI-mediated writing environments remains fundamentally social and mediated, although the mediating tools have changed. Vygotsky (1978) conceptualized learning as developing

through interaction, scaffolding, and guided participation. In the present study, AI functioned as one of several mediational tools, alongside instructor feedback, peer interaction, reflection, and revision, that supported students' writing development. However, findings consistently emphasized that meaningful learning depended not on AI-generated output itself but on students' ability to interpret, evaluate, and strategically incorporate AI-generated suggestions. This finding supports recent work positioning generative AI as a reflective learning partner rather than a replacement for human judgment (Tian & Louw, 2026). Similarly, Barrot (2024) argues that AI should be integrated within pedagogically guided writing instruction that continues to prioritize critical thinking, revision, and learner agency rather than automated text production.

One of the most significant contributions of this study is its challenge to conventional product-based assessment. Traditional assessment practices have generally assumed that final written products provide adequate evidence of student learning. However, findings repeatedly described how AI-assisted drafting, editing, paraphrasing, and language revision blurred the relationship between final text quality and actual learning. These findings suggest that assessment validity increasingly depends on documenting the writing process rather than evaluating only polished outcomes. This interpretation aligns with Tian and Louw's (2026) argument that assessment should shift from evaluating products toward examining learners' reasoning, reflective dialogue, and decision-making processes. It also complements Seo's (2024) findings that students' prompts, reflections, and revision activities reveal substantially richer evidence of learning than final essays alone. Together, these studies reinforce the present findings that drafting histories, revision logs, reflective commentaries, and AI-use documentation provide more meaningful evidence of writing development than isolated final submissions.

The findings also contribute to ongoing discussions of authorship in AI-assisted writing. Rather than conceptualizing authorship as a binary distinction between human-generated and AI-generated text, instructors described authorship as emerging through students' decisions to accept, reject, revise, verify, and explain AI-generated suggestions. This perspective extends recent scholarship arguing that learner agency remains central in AI-supported writing (Dong, 2024; Pack & Maloney, 2024). Students demonstrated authorship not simply by producing original text but by exercising informed judgment throughout the writing process. Consequently, authorship became visible through documented thinking, revision choices, and reflective justification rather than through attempts to determine whether AI had been used. This finding represents an important conceptual shift from detecting AI toward understanding how students engage with AI during learning.

Another important contribution concerns the role of reflection and metacognition in AI-mediated writing. Reflection activities enabled instructors to examine students' reasoning, evaluate how they interpreted feedback, and understand why particular

revisions were made. Reflection therefore functioned as both a learning strategy and an assessment tool. These findings align with feedback literacy research emphasizing that meaningful learning occurs when students actively interpret, evaluate, and act upon feedback (Carless & Boud, 2018). Likewise, Dong (2024) argues that AI-assisted writing requires ethical engagement in addition to cognitive and behavioral engagement, requiring learners to critically evaluate AI-generated feedback rather than accept it uncritically. The present findings suggest that structured reflection creates opportunities for students to demonstrate this ethical and metacognitive engagement while simultaneously strengthening AI literacy.

The findings further suggest that process-oriented assessment provides a constructive alternative to AI surveillance and detection. Participants consistently questioned the effectiveness of AI-detection tools and instead emphasized assessment practices that encourage openness about AI use. AI-use disclosures, revision logs, multiple drafts, source-verification activities, and reflective commentaries shifted assessment from identifying misconduct toward documenting responsible learning. This approach resonates with recent scholarship advocating transparency, human oversight, and ethical AI integration instead of prohibition or punishment (Al-khresheh, 2024; Kostka & McMartin-Miller, 2026; Pitychoutis & Spathopoulou, 2026). Rather than discouraging AI use altogether, these practices help students develop the critical evaluation, verification, and decision-making skills required for responsible participation in AI-mediated academic communities.

Finally, the findings support a broader reconceptualization of academic integrity in higher education. Traditionally, academic integrity has been associated primarily with preventing plagiarism and unauthorized assistance. However, participants described integrity as a process demonstrated through transparency, accountability, documentation, and evidence of learning. Within this perspective, responsible AI use is evaluated not by whether AI was involved but by how students engaged with AI throughout the writing process. This reconceptualization shifts the focus from policing technology to supporting ethical scholarly practices and aligns with emerging calls to redesign assessment around learning documentation rather than AI detection.

In general, the findings suggest that process-oriented assessment represents more than a collection of classroom strategies; it offers a conceptual framework for rethinking assessment, authorship, and academic integrity in AI-mediated writing. By emphasizing visible learning, reflective decision-making, and responsible AI engagement, this approach provides a more valid and educationally meaningful foundation for assessing multilingual writers in contemporary higher education.

Recommendations

The findings suggest that assessment in AI-mediated academic writing should move beyond evaluating final essays to documenting how learning develops throughout the writing process. Instructors can redesign rubrics to include evidence such as

brainstorming, planning, multiple drafts, revision logs, reflective commentaries, writing conferences, and transparent AI-use disclosures. Such evidence provides a more complete picture of students' reasoning, revision decisions, and writing development than the finished product alone. Clear course-level AI policies should define appropriate uses of AI while encouraging disclosure rather than concealment. Equally important, writing instruction should explicitly develop students' critical AI literacy, source verification, feedback literacy, and reflective decision-making so that AI supports, rather than replaces, independent learning and academic judgment.

Implications

The findings suggest that AI-mediated writing requires a broader rethinking of assessment rather than minor adjustments to existing practices. For TESOL practitioners, the study indicates that evaluating final essays alone may no longer provide sufficient evidence of students' learning, authorship, or critical engagement. Process-oriented assessment, including multiple drafts, revision histories, reflective commentaries, writing conferences, and AI-use disclosures, offers instructors a more complete understanding of how students construct knowledge, respond to feedback, and make writing decisions over time. For curriculum designers, these findings point to the need to embed AI literacy, feedback literacy, ethical source use, and reflective practice into academic writing curricula instead of treating them as supplementary skills. At the institutional level, the findings suggest that academic integrity policies may be more effective when they promote transparency, documentation, and responsible AI engagement rather than relying primarily on AI detection technologies. Such an approach better reflects how writing is evolving in higher education while preserving human judgment, learner accountability, and meaningful evidence of learning.

Limitations and Future Research

This study should be interpreted in light of several limitations. As a qualitative investigation conducted within a specific academic writing context, the findings are intended to provide contextual understanding rather than broad generalization. Because AI technologies and instructional practices continue to evolve, assessment approaches may also change over time. Future research could examine the long-term effects of process-oriented assessment on students' writing development, AI literacy, and ethical source use. Comparative studies across disciplines, educational levels, and institutional contexts would further clarify the transferability of these findings. Research evaluating specific assessment practices, such as revision logs, reflective writing, and AI-use disclosures, may also provide practical guidance for instructors redesigning assessment in AI-mediated learning environments.

Conclusion

This study argues that generative AI requires a fundamental reconsideration of how academic writing is assessed rather than simply new methods for detecting AI use. The findings suggest that process-oriented assessment provides a more valid

representation of learning by documenting students' drafting, revision, reflection, and decision-making throughout the writing process. Guided by Sociocultural Theory, the study reconceptualizes assessment as evidence of learning rather than evaluation of final products alone. By positioning transparency, responsible AI engagement, and human judgment at the center of assessment, the proposed framework offers a practical direction for supporting academic integrity, preserving student authorship, and assessing multilingual writers in AI-mediated higher education classrooms.

ORCID

 <https://orcid.org/0000-0001-5548-7141>

Publisher's Note

The claims, arguments, and counter-arguments made in this article are exclusively those of the contributing authors. Hence, they do not necessarily represent the viewpoints of the authors' affiliated institutions, or EUROKD as the publisher, the editors and the reviewers of the article.

Acknowledgements

The author is grateful to the editor and the anonymous reviewers for their careful evaluation of the manuscript and for their thoughtful suggestions. Their constructive comments were invaluable in strengthening the quality and clarity of this work.

Funding

No external funding was received for this research.

CRedit Authorship Contribution Statement

Laila Noor: Conceptualization, Methodology, Investigation, Data Curation, Formal Analysis, Visualization, Writing, Original Draft, Writing, Review & Editing

Generative AI Use Disclosure Statement

ChatGPT (OpenAI) was used solely to assist with language refinement, including improvements to grammar, clarity, organization, and readability, as well as for figure creation. It was not used for the conceptualization of the study, literature synthesis, research design, data analysis, interpretation of findings, or the development of scholarly arguments. The author assumes full responsibility for the accuracy, integrity, and intellectual content of the manuscript.

Ethics Declarations

World Medical Association (WMA) Declaration of Helsinki–Ethical Principles for Medical Research Involving Human Participants

This study involved classroom-based educational research with advanced English learners and writing instructors in a U.S. higher education context. Participation and the use of classroom-generated data were conducted with appropriate attention to participant privacy, confidentiality, and ethical research practices. All data used for

research purposes were anonymized to protect participants' identities. The study did not involve medical research, clinical intervention, or experimentation and therefore did not involve the types of medical procedures addressed by the World Medical Association Declaration of Helsinki.

Competing Interests

The author declares that there are no conflicts of interest related to this study.

Data Availability

The data underlying this study are available from the corresponding author upon reasonable request.

References

- Al-khresheh, M. H. (2024). The future of artificial intelligence in English language teaching: Pros and cons of ChatGPT implementation through a systematic review. *Language Teaching Research Quarterly*, 43, 54–80. <https://doi.org/10.32038/ltrq.2024.43.04>
- Barrot, J. S. (2024). Leveraging ChatGPT in the writing classrooms: Theoretical and practical insights. *Language Teaching Research Quarterly*, 43, 43–53. <https://doi.org/10.32038/ltrq.2024.43.03>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp0630a>
- 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>
- Braun, V., & Clarke, V. (2021a). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology*, 18(3), 328–352. <https://doi.org/10.1080/14780887.2020.1769238>
- Braun, V., & Clarke, V. (2021b). *Thematic Analysis: A Practical Guide*. SAGE.
- Braun, V., & Clarke, V. (2023). Toward good practice in thematic analysis: Avoiding common problems and be(com)ing a knowing researcher. *International Journal of Transgender Health*, 24(1), 1–6. <https://doi.org/10.1080/26895269.2022.2129597>
- Carless, D., & Boud, D. (2018). The development of student feedback literacy: enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, 43(8), 1315–1325. <https://doi.org/10.1080/02602938.2018.1463354>
- Dong, L. (2026). When machines teach writers: Reflecting on GenAI's role in giving corrective feedback on linguistic CAF facets. *Feedback Research in Second Language*, 4, 28–39. <https://doi.org/10.32038/frsl.2026.04.04>
- Dong, L. (2024). ChatGPT in language writing education: Reflections and a research agenda for a ChatGPT feedback engagement framework. *Language Teaching Research Quarterly*, 43, 121–131. <https://doi.org/10.32038/ltrq.2024.43.07>
- Escalante, J., Pack, A., & Barrett, A. (2023). AI-generated feedback on writing: Insights into efficacy and ENL student preference. *International Journal of Educational Technology in Higher Education*, 20(57), 1–15. <https://doi.org/10.1186/s41239-023-00425-2>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kostka, I., & McMartin-Miller, C. (2026). Reflections on using GenAI for multilingual writing feedback: Implications for practice. *Feedback Research in Second Language*, 4, 16–27. <https://doi.org/10.32038/frsl.2026.04.03>
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage Publications.
- Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study. *Smart Learning Environments*, 11(9), 1–20. <https://doi.org/10.1186/s40561-024-00295-9>
- Mollick, E. (2024). *Co-intelligence: Living and working with AI*. Portfolio.

- Noor, L., & Islam, A. S. M. K. (2025). Enhancing English language skills with ChatGPT: A focus on writing, reading comprehension, and vocabulary development. *SSTESOL Journal*, 17(1), 20–30. <https://pub.marq.com/38f59a01-eabe-4f2d-8279-5ed78debfc9/#sUwOQlCNZWa7>
- Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, 16, 1–13. <https://doi.org/10.1177/1609406917733847>
- Pack, A., & Maloney, J. (2024). Using artificial intelligence in TESOL: Some ethical and pedagogical considerations. *TESOL Quarterly*, 58(2), 1007–1018. <https://doi.org/10.1002/tesq.3320>
- Pitychoutis, K. M., & Spathopoulou, F. (2026). AI-mediated written corrective feedback in L2 writing: What changes, what doesn't, and what we still don't know. *Feedback Research in Second Language*, 4, 6–15. <https://doi.org/10.32038/frsl.2026.04.02>
- Sandelowski, M. (2000). Whatever happened to qualitative description? *Research in Nursing & Health*, 23(4), 334–340. [https://doi.org/10.1002/1098-240X\(200008\)23:4<334::AID-NUR9>3.0.CO;2-G](https://doi.org/10.1002/1098-240X(200008)23:4<334::AID-NUR9>3.0.CO;2-G)
- Sandelowski, M. (2010). What's in a name? Qualitative description revisited. *Research in Nursing & Health*, 33(1), 77–84. <https://doi.org/10.1002/nur.20362>
- Schwartz, H. L., & Diliberti, M. K. (2026). *More students use AI for homework, and more believe it harms critical thinking: Selected findings from the American Youth Panel (RR-A4742-1)*. RAND Corporation. https://www.rand.org/pubs/research_reports/RRA4742-1.html
- Selim, A. S. M. (2024). The transformative impact of AI-powered tools on academic writing: Perspectives of EFL university students. *International Journal of English Linguistics*, 14(1), 14–26. <https://doi.org/10.5539/ijel.v14n1p14>
- Seo, J.-Y. (2024). Exploring the educational potential of ChatGPT: AI-assisted narrative writing for EFL college students. *Language Teaching Research Quarterly*, 43, 1–21. <https://doi.org/10.32038/ltrq.2024.43.01>
- Starominski-Uehara, M. (2026, March 26). *GenAI and creeping cognitive displacement syndrome*. EDUCAUSE Review. <https://er.educause.edu/articles/2026/3/genai-and-creeping-cognitive-displacement-syndrome>
- Subedi, R., & Nyamasvisva, T. E. (2024). A review of the influence of artificial intelligence in academic writing. *Journal of Computer Science and Information Technology*, 2(1), 71–80. <https://doi.org/10.70248/jcsit.v2i1.1134>
- Tian, W., & Louw, S. (2026). DeepThink-mediated language assessment literacy: EFL teachers' practices and insights from China and Cambodia. *Language Testing in Focus: An International Journal*, 13, 77–88. <https://doi.org/10.32038/ltf.2026.13.05>
- Veldhuis, A., Lo, P. Y., Kenny, S., & Antle, A. N. (2025). Critical artificial intelligence literacy: A scoping review and framework synthesis. *International Journal of Child-Computer Interaction*, 43, 100708. <https://doi.org/10.1016/j.ijcci.2024.100708>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Zhai, X. (2022). ChatGPT user experience: Implications for education. *Educational Technology Research and Development*. <http://dx.doi.org/10.2139/ssrn.4312418>