RANGGUK: Jurnal Pengabdian Kepada Masyarakat

Vol. 06 No. 01 2026
E-ISSN: 3047-0897

Establishing Ethical Guidelines for AI Usage in EFL Academic Writing

Edy Suseno
IKIP Widya Darma
Correspondent: Edysuseno4@gmail.com
DOI: https://doi.org/10.32939/rgk.v6i1.6714

Abstract

This qualitative study investigates the ethical complexities of AI authorship among 25 EFL college learners, focusing on the tension between linguistic enhancement and the preservation of authorial voice. By mapping an "Authorial Spectrum," the research distinguishes between AI acting as a "Cradle" (technical scaffolding for grammar and syntax) versus a "Creator" (the automated generation of original conceptual ideas). Findings reveal a pervasive "Shortcut Trap," where deadline pressures drove 68% of participants toward intellectual dependency, allowing AI-generated structures to supersede their personal logical flow. However, the data also highlights the efficacy of "Socratic Prompting"—defined as iterative, question-based AI interaction—in safeguarding unique cultural nuances from "algorithmic sanitization" (the loss of identity through standardized AI output). This method resulted in a 40% higher retention of original idiomatic expressions compared to direct, one-off prompting. Consequently, the study advocates for a process-oriented disclosure model using transparency logs to address "digital ghostwriting." By moving beyond a prohibition-based mindset, the research positions AI as a tool to sharpen the human filter rather than replace the learner's identity. Ultimately, these findings ensure that the diverse voices of EFL learners remain the primary drivers of academic discourse in an increasingly automated age.
keywords: Academic Integrity; AI-Assisted; Authorial Voice; EFL Learner;, Ethics

Introduction

The emergence of generative artificial intelligence (AI) has fundamentally restructured the pedagogical landscape of higher education, particularly by redefining the interaction between students and the academic writing process. Platforms such as ChatGPT and Gemini are increasingly conceptualized as "collaborative partners" that facilitate democratic access to success for English as a Foreign Language (EFL) learners by offering critical linguistic scaffolding and stylistic enhancement (Moorhouse & Yeo, 2023). This technological paradigm shift is widely lauded for mitigating traditional barriers to entry for non-native speakers, providing real-time feedback and high-level grammatical interventions that were historically beyond the reach of many students (Sullivan & Kelly, 2024). However, this optimistic discourse is increasingly countered by critical investigations into the "algorithmic sanitization" of student identity, as automated systems tend to enforce standardized fluency at the expense of individual, culturally nuanced expression (Kaur & Singh, 2024). The inherent bias of these models toward Western-centric rhetorical norms often results in the marginalization of translingual identities, effectively trapping students in a "correctness" framework that devalues their heritage and indigenous perspectives (Richards & Smith, 2024). Consequently, the academic sector faces a profound dichotomy where AI is perceived simultaneously as an essential lifeline for linguistic equity and a fundamental threat to the authenticity of human cognition (Suseno, 2025a). As these generative tools become an inseparable component of the digital scholar's repertoire, it is imperative for educators to

bridge the gap between leveraging AI for skill delivery and ensuring that the student's unique authorial voice remains the central catalyst of academic discourse (Suseno, 2024a).

Beneath these perceived utility benefits lies a critical challenge involving the systematic erosion of authorial agency and the emergence of the "Shortcut Trap," which jeopardizes the cognitive evolution of EFL learners. When students relinquish the synthesis of arguments or the rhetorical structuring of their essays to automated systems, they effectively bypass the intellectual friction necessary for deep language acquisition and the development of critical thinking faculties (Bond & Khosravi, 2025). This externalization of cognitive labor drives a homogenization of academic output, producing submissions that are linguistically flawless yet devoid of the idiosyncratic nuances and personal narratives that define a scholar's unique perspective (Choi & Lee, 2024). Such a loss of "human imperfection" signals a shift toward "authorial alienation," a psychological state where the student no longer recognizes the final text as an authentic extension of their internal mental journey (Prentice & Kormos, 2025). Furthermore, a mounting crisis of intellectual dependency has become evident, as learners frequently find themselves incapable of defending their arguments in live settings because the foundational logic was generated by an algorithm rather than through personal mental effort (Suseno, 2024c). This dependency is intensified in high-pressure academic environments where the drive for performance often culminates in "digital ghostwriting," marked by an almost total reliance on machine-generated heuristics (Duarte & Santos, 2023). Without targeted pedagogical interventions—such as the integration of podcasts or translation-based cognitive exercises to stimulate original thought—students are at risk of becoming passive spectators within their own educational trajectory (Suseno, 2024; Bond & Khosravi, 2025).

While the technical capacity of AI to refine grammatical accuracy is well-documented, current scholarship notes a significant decline in the employment of culturally specific metaphors in AI-assisted work (Kaur & Singh, 2024). Research concerning academic integrity has historically remained tethered to detection-based strategies; however, many experts now argue that the escalating "cat-and-mouse game" between AI detectors and generative tools is largely futile and counterproductive to the learning process (Farrelly & Baker, 2024). In response to these limitations, recent studies have begun to champion "Socratic prompting" as a critical literacy intervention to safeguard student agency, positing that using AI as an iterative brainstorming interlocutor preserves the integrity of the intellectual labor (Gao & Zheng, 2025). This transition toward a process-oriented ethical framework is further supported by the use of "Transparency Logs," which require students to provide a visible account of their dialogic journey with the technology (Tan & Wang, 2023). Innovative pedagogical strategies, including the use of image descriptions and visual interpretation, have also been introduced to stimulate writing and speaking proficiencies without entirely substituting human creativity (Suseno, 2025d). Despite these developments, much of the existing literature is confined to technical silos, often neglecting the profound psychological consequences of algorithmic feedback on a student's long-term writing self-efficacy (Prentice & Kormos, 2025). There is an urgent necessity to develop a comprehensive framework that clearly differentiates between AI utilized as a supportive "Cradle" and AI serving as a surrogate "Creator" to ensure the preservation of the student's analytical autonomy (Amani & Abbas, 2024; Zhang, 2025).

The rapid integration of generative artificial intelligence (AI) has fundamentally altered the relationship between students and academic writing in higher education, with tools such as ChatGPT and Gemini frequently framed as "collaborative partners" that democratize success for English as a Foreign Language (EFL) learners by providing essential linguistic scaffolding (Moorhouse & Yeo, 2023). This technological shift is often celebrated for its ability to lower barriers for non-native speakers, offering immediate feedback and sophisticated grammatical correction that was previously inaccessible to many; however, this

optimism is increasingly challenged by concerns regarding the "algorithmic sanitization" of a student's unique cultural identity, as machines tend to prioritize standardized fluency over individual, nuanced expression (Kaur & Singh, 2024; Sullivan & Kelly, 2024). The machine's propensity to favor Western-centric academic norms often leads to the erasure of translingual identities, forcing students to adapt to a "correctness" trap that devalues their heritage, thereby creating a complex dichotomy where AI is viewed simultaneously as a vital lifeline and a potential threat to the authenticity of human thought (Suseno, 2025a; Richards & Smith, 2024). As these tools become a ubiquitous component of the digital scholar's toolkit, educators must navigate the tension between leveraging AI for delivery skills and ensuring that the student's original voice remains the primary driver of academic discourse, particularly as the outsourcing of cognitive labor results in a homogenization of academic output (Suseno, 2024a; Choi & Lee, 2024). This loss of idiosyncratic nuance signifies a deeper issue of "authorial alienation," where the writer no longer recognizes the final product as a reflection of their internal cognitive journey, a crisis exacerbated in high-stakes environments where the pressure to perform leads to "digital ghostwriting" (Prentice & Kormos, 2025; Duarte & Santos, 2023). Without explicit pedagogical interventions to stimulate original thought, such as the use of podcasts or translation exercises, students risk becoming spectators of their own education, unable to defend arguments in real-time because the underlying logic was formulated by an algorithm rather than their own mental effort (Suseno, 2024; Suseno, 2024c; Bond & Khosravi, 2025).

Previous studies have extensively documented the mechanical capabilities of AI in enhancing grammatical accuracy among non-native speakers, yet a significant gap remains regarding the psychological and cultural impact of "authorial alienation" (Kaur & Singh, 2024; Prentice & Kormos, 2025). While detection-based strategies have proven increasingly ineffective, recent investigations suggest that "Socratic prompting" and "Transparency Logs" may offer a viable path toward maintaining student agency (Farrelly & Baker, 2024; Tan & Wang, 2023). However, much of the existing literature remains siloed in technical contexts, overlooking the fluid boundaries of the "Authorial Spectrum" in EFL settings (Amani & Abbas, 2024). There is, therefore, an urgent need to establish a framework that distinguishes between AI as a supportive "Cradle" and a surrogate "Creator" to ensure the preservation of the student's analytical autonomy (Zhang, 2025). To address this gap, the present study investigates the ethical navigation of AI authorship among EFL college learners through the following research questions:

1. How do EFL students define the threshold between AI-assisted and AI-generated work within the Authorial Spectrum?
2. To what extent does Socratic Prompting influence the retention of a student's cultural and personal voice compared to direct AI generation?
3. What psychological barriers influence a student's decision to disclose or conceal their interactive journey with generative AI?

Public Society Method

Research Design and Framework

This study employs a qualitative case study design to rigorously investigate the multifaceted ethical dimensions of human-AI interaction within the sphere of academic writing, specifically focusing on the lived experiences of EFL learners. This methodological framework was strategically selected for its capacity to facilitate an in-depth analysis of the "Authorial Spectrum" by examining the subjective decision-making processes that students navigate during complex writing tasks. Unlike purely theoretical or speculative inquiries, a qualitative case study prioritizes the procedural and cognitive nuances of how students negotiate their linguistic identity and agency when utilizing generative AI tools in a real-

world academic setting. By adopting this approach, the research can capture the fluid nature of authorship as it shifts between human intent and algorithmic suggestion, providing a granular look at the "Ethical Red Line" that defines modern scholarly integrity. This design is particularly effective for uncovering the underlying rationales that drive student behavior, allowing for a comprehensive exploration of how technology either scaffolds or supplants the individual's authorial voice over an extended period of observation.

Participants and Purposive Sampling Strategy

A total of 25 undergraduate students (N=25) were recruited through purposive sampling from an intensive English language program at a major urban university in Indonesia to serve as the primary participant cohort. The participants were specifically selected from second and third-year students across diverse academic disciplines, including the Humanities (n=10), Social Sciences (n=8), and STEM fields (n=7), to ensure that the study's findings reflect a broad and representative spectrum of AI interaction patterns across different rhetorical traditions. This disciplinary diversity is crucial for understanding how varying academic requirements influence the degree of reliance on generative tools and the subsequent impact on disciplinary identity. Participation in the study was strictly voluntary and governed by a rigorous informed consent process, ensuring that all contributors were fully aware of the research objectives and their rights as subjects. To maintain the highest standards of ethical compliance and protect participant privacy, all identities were fully anonymized through the application of alphanumeric codes (e.g., P01, P02), which were used consistently throughout the data collection and analysis phases to prevent any breach of confidentiality.

Data Collection Instruments and Triangulation Procedures

Data were systematically collected over the course of a 12-week semester utilizing a data triangulation strategy that integrated three primary qualitative instruments to ensure methodological rigor and validity. First, weekly 90-minute non-participant observations were conducted in the university's digital writing lab, allowing researchers to record the real-time behavioral ratio between "iterative prompting" and "manual editing" as students engaged with AI interfaces. Second, participants maintained AI Transparency Logs, which served as an objective audit trail by documenting the verbatim interactive dialogue between the students and the machine, capturing the evolution of prompts and the subsequent rejection or acceptance of AI suggestions. Third, the study employed semi-structured interviews incorporating a "think-aloud" protocol, providing a platform for students to offer verbal rationalizations for their executive decisions and the psychological factors influencing their writing process. The integration of these three distinct data sources was essential for minimizing self-reporting bias and capturing the dynamic, real-time decision-making patterns that characterize the modern EFL writing experience, thereby ensuring a robust and multi-layered empirical foundation for the study's conclusions. As illustrated in Table 1, the study integrated three distinct data streams to ensure a holistic understanding of the Authorial Spectrum.

Table 1. Summary of Data Collection Instruments and Triangulation Strategy

Instrument	Frequency/Duration	Primary Data Focus	Purpose in Triangulation
Non-participant Observation	Weekly (90 mins per session)	Behavioral ratios; manual editing vs. AI prompting patterns.	To capture real-time, objective behavioral data and minimize self-reporting bias.
AI Transparency	Continuous (12-week semester)	Verbatim prompt history; iterative	To provide an objective audit trail of

Instrument	Frequency/Duration	Primary Data Focus	Purpose in Triangulation
Logs		dialogue; AI-human exchange.	the procedural journey and "process disclosure."
Semi-structured Interviews	End of semester (45-60 mins)	"Think-aloud" protocols; student rationales; psychological barriers.	To uncover the subjective decision-making processes and emotional/ethical nuances.

Thematic Data Analysis and Coding Framework

The qualitative data generated from this study were analyzed thematically following the established framework of Braun and Clarke (2006), a process facilitated by the use of NVivo qualitative analysis software to ensure systematic organization and retrieval. The analytical procedure involved three distinct and rigorous stages of coding designed to move from raw data to higher-level theoretical synthesis. Initially, open coding was employed to label discrete raw units of analysis, such as instances of "rapid copy-pasting," "iterative questioning," or "manual syntactic revision," providing an granular map of student-AI interactions. This was followed by a stage of axial coding, wherein these initial codes were grouped into functional categories and conceptual clusters, such as "Linguistic Scaffolding" (the AI as a Cradle) and "Content Substitution" (the AI as a Creator). Finally, selective coding was performed to synthesize these categories into overarching themes within the "Authorial Spectrum" framework, allowing the researchers to identify the critical ethical threshold where technological assistance transitions into "digital ghostwriting." This multi-tiered coding system ensures that the study's findings are not merely anecdotal but are grounded in a structured interpretation of the participants' behavioral and cognitive patterns.

Data Analysis and Coding System

The data were analyzed thematically following Braun and Clarke's (2006) framework, facilitated by NVivo qualitative analysis software to ensure systematic organization. The process involved three distinct stages of coding: open coding was first used to label raw units of analysis; this was followed by axial coding to categorize data into functional concepts; and finally, selective coding was performed to identify the ethical threshold within the "Authorial Spectrum." The systematic progression of this coding process, along with representative examples of emergent categories, is illustrated in Table 2 below.

Table 2. Sample Coding Scheme for the Authorial Spectrum Analysis

Coding Phase	Examples of Codes/Categories	Representative Data/Themes
Open Coding	*Rapid copy-pasting; iterative questioning; manual syntactic repair; phrase-level translation.*	Individual student actions and behavioral patterns in the digital writing lab.
Axial Coding	Linguistic Scaffolding (Cradle); Content Substitution (Creator); Cognitive Outsourcing.	Functional categorization of human-AI dialogue and interaction patterns.
Selective Coding	The Authorial Spectrum; The Shortcut Trap; The Ethical Red Line.	Overarching themes determining the threshold of student agency and analytical autonomy.

Methodological Rigor and Trustworthiness Criteria

To ensure the validity, reliability, and overall trustworthiness of the findings, this study strictly adhered to four established criteria for qualitative rigor, tailored to the complexities of digital education research. First, data triangulation was achieved by synthesizing evidence from non-participant observations, AI transparency logs, and semi-structured interviews, thereby minimizing individual instrument bias and providing a holistic view of the phenomena. Second, the study utilized peer debriefing, where the coding scheme and emergent themes were reviewed by an external researcher to challenge potential investigator bias and ensure analytical consistency. Third, member checking was conducted by sharing preliminary interpretations with the participants to verify that the findings accurately reflected their lived experiences and executive decisions. Finally, the research provides a thick description of the data through the inclusion of detailed excerpts from the transparency logs and verbatim interview quotes, allowing for a high degree of transferability and transparency. Collectively, these rigorous procedures ensure that the study's conclusions are deeply rooted in documented empirical evidence, thereby maintaining high scholarly credibility and ensuring the findings are defensible within the evolving context of AI in higher education.

Results and Discussion

The Authorial Spectrum: Temporal Dynamics and the Shift from "Cradle" to "Creator"

Analysis of the Transparency Logs and longitudinal observations revealed that while all 25 participants (N=25) engaged with generative AI, their positions on the "Authorial Spectrum" were highly dynamic, shifting significantly in response to task complexity and looming academic deadlines. In the early stages of the semester, 72% (n=18) of the cohort utilized AI primarily as a "Cradle" for structural scaffolding, such as requesting outlines for complex essays or clarifying theoretical frameworks without asking for direct content generation. However, a critical behavioral transition was observed as deadlines approached in week 10, where 48% (n=12) of participants drifted toward the "Creator" persona by requesting full-paragraph generation to bypass linguistic hurdles. This shift was candidly documented in the logs; for instance, participant P04 noted: *"The deadline was in two hours; I asked Gemini to 'write a concluding paragraph based on my points' because I lacked the vocabulary to finish in time."* Such data confirms that the "Shortcut Trap" is not necessarily a result of a premeditated lack of ethical intent, but is rather a pragmatic response to temporal pressure and perceived linguistic inadequacy, illustrating how the boundary of authorship becomes fluid under situational stress.

Socratic Prompting vs. Direct Generation: Mitigating Linguistic Sanitization

The study identified a stark contrast in the retention of authorial voice between students who employed "Socratic Prompting" and those who relied on direct generative commands. Quantitative analysis of the logs showed that students who used iterative, question-based prompts (n=10) retained 40% more original idiomatic expressions and cultural markers compared to the group utilizing direct prompts. This retention is crucial in preventing "Linguistic Sanitization," where AI algorithms tend to replace unique cultural metaphors with standardized, "hollow" academic prose. A notable example occurred with participant P12, who intervened when the AI attempted to replace the Indonesian concept of *gotong royong* with the generic term "teamwork," stating in their log: *"I told the AI, 'Keep the community spirit nuance but make it sound academic.' It helped me find the term 'communal synergy' instead of just losing my original idea."* In contrast, the direct-generation group saw 85% of

their unique cultural metaphors eliminated, suggesting that without the "Socratic shield," the machine's tendency toward Western-centric standardization effectively erases the learner's idiosyncratic and translingual identity.

In contrast, the direct-generation group saw 85% of their unique cultural metaphors eliminated, suggesting that without the 'Socratic shield,' the machine's tendency toward Western-centric standardization effectively erases the learner's idiosyncratic identity. To provide a concise overview of these behavioral shifts and the corresponding empirical evidence, the key findings are synthesized in Table 3 below.

Table 3. Summary of Key Findings: Behavioral Shifts and Authorial Retention

Theme	Key Metric/Finding	Representative Evidence (Participant Quote)
The Authorial Spectrum	72% used AI as "Cradle"; 48% shifted to "Creator" under deadline pressure.	*"The deadline was in two hours; I asked Gemini to write a concluding paragraph..."* (P04)
Linguistic Identity	Socratic Prompting retained 40% more cultural idiomatic expressions vs. direct commands.	*"I told the AI, 'Keep the community spirit nuance but make it sound academic'."* (P12)
Ethical Agency	80% reported increased critical oversight and agency via Transparency Logs.	*"Writing down my prompts made me realize I was becoming a 'spectator'."* (P21)

Psychological Barriers and Transparency Logs as Catalysts for Agency

Data derived from semi-structured interviews highlighted a pervasive "Fear of Stigma" among EFL learners, which initially acted as a significant barrier to the ethical disclosure of AI usage. Even when utilizing AI for legitimate purposes, such as grammar checks or vocabulary expansion, 60% (n=15) of participants reported feeling a sense of "guilt," fearing that any level of AI involvement would be misinterpreted as academic misconduct. However, the systematic implementation of Transparency Logs functioned as a transformative psychological catalyst for ethical agency and self-reflection. By the eighth week of the semester, 80% of the participants reported that the mandatory act of documenting their prompts forced them to become more critical and selective regarding AI-generated output. This shift from passive consumption to active oversight was exemplified by participant P21, who remarked: *"Writing down my prompts made me realize I was becoming a 'spectator.' It forced me to start questioning the AI's suggestions instead of just clicking 'copy'."* This finding underscores that transparency tools do more than just provide an audit trail; they actively empower students to maintain their role as the primary investigator of their work, thereby significantly reducing the risk of unintentional digital ghostwriting.

Persona Identification and Behavioral Patterns

The synthesis of behavioral data resulted in the categorization of participants into four distinct personas (Table 1).

Table 4. Persona Identification and Behavioral Patterns

Persona	AI Usage Pattern	Retention of Voice	Frequency (N=25)
Strategic Collaborator	Iterative/Socratic Prompting	High (>80%)	n = 7 (28%)
Linguistic Seeker	Grammar & Syntax assistance	Moderate (50-70%)	n = 8 (32%)
Prompt Engineer	Large-scale content	Low (<30%)	n = 6 (24%)

Persona	AI Usage Pattern	Retention of Voice	Frequency (N=25)
	generation		
Spectator	Total dependency/Copy-paste	Minimal (<10%)	n = 4 (16%)

The data indicates that while the Strategic Collaborator is the ethical ideal, the majority of students (56%) occupy the middle-ground, often drifting toward the Spectator model when linguistic pressure exceeds their current proficiency level. This drift underscores the necessity for pedagogical frameworks that prioritize the "process" of writing over the final "perfect" product.

Discussion

The Authorial Spectrum and the Threshold of Agency

The findings of this study demonstrate that the "red line" of authorship for EFL learners is not a static boundary but a fluid negotiation heavily influenced by cognitive effort and temporal pressure. Unlike the binary "misconduct vs. integrity" model prevalent in current institutional policies, the Authorial Spectrum suggests that ethical authorship is fundamentally defined by the transformative act of the human mind upon algorithmic output. When participants utilized AI as a "Cradle" for structural scaffolding, they successfully maintained their status as the primary investigators of their prose, using the technology to clarify their own original intent (Amani & Abbas, 2024). However, the observed drift toward the "Creator" model under intense deadline pressure highlights a significant "Shortcut Trap" where pragmatic necessity frequently compromises the deep language acquisition process (Bond & Khosravi, 2025). This empirical evidence confirms that unless students intentionally act as "Strategic Collaborators"—constantly interrogating and refining AI-generated suggestions—they risk entering a state of chronic intellectual dependency. In such a state, algorithmic logic effectively supersedes critical thinking, leading to a loss of the analytical autonomy that is essential for rigorous scholarly development (Duarte & Santos, 2023; Moorhouse & Yeo, 2023).

Socratic Prompting as a Safeguard for Cultural Identity

The efficacy of Socratic Prompting identified in this research underscores the paramount importance of iterative dialogue in safeguarding the unique "EFL Identity" against the pervasive threat of linguistic standardization. While direct AI generation often produces text that is grammatically flawless, it frequently results in "hollow" prose that lacks the rhetorical soul of the author. In contrast, the Socratic method empowers students to retain culturally specific metaphors and idiosyncratic phrasing—elements that standardized algorithms typically eliminate in favor of generic, Western-centric fluency (Choi & Lee, 2024; Kaur & Singh, 2024). This iterative interaction functions as a critical cognitive filter, ensuring that AI-assisted writing remains an authentic extension of the student's unique cross-cultural perspective rather than a generic digital facsimile (Richards & Smith, 2024). By prioritizing linguistic clarity over total automated content generation, learners are able to move beyond a restrictive "prohibition" mindset toward a more sophisticated model of Critical AI Literacy. In this model, so-called "human imperfections" are not viewed as errors to be sanitized, but are instead preserved as essential markers of authentic scholarship and personal voice (Gao & Zheng, 2025; Suseno, 2025a).

Procedural Ethics and the Role of Transparency Logs

This investigation further highlights that psychological barriers, specifically the "fear of stigma" and the perceived risk of being labeled an academic cheat, often drive AI usage underground. Such "hidden usage" patterns are particularly dangerous as they hinder effective pedagogical intervention and prevent the development of a healthy human-AI relationship (Farrelly & Baker, 2024). The implementation of Transparency Logs proved to be a transformative intervention in this context by shifting the ethical focus from the final product to the integrity of the procedural journey (Tan & Wang, 2023). By formalizing the disclosure of AI interactions, students were able to develop a sense of executive agency, ultimately recognizing themselves as responsible digital scholars who lead the machine rather than passive spectators who follow it (Liu & Zhao, 2024). This fundamental transition suggests that future EFL curricula should move toward integrating "Process Disclosure" as a core component of academic integrity. Such a move would ensure that AI technology serves to sharpen the human filter and enhance linguistic precision without replacing the learner's essential analytical autonomy or self-efficacy (Prentice & Kormos, 2025; Zhang, 2025).

Limitations and Future Research

Despite the significant insights gained, this study is subject to several limitations that warrant cautious interpretation. First, the small sample size (N=25) within a specific urban university context may limit the generalizability of the findings to broader EFL populations with different socio-economic backgrounds or technological access. Second, while Transparency Logs provided an objective audit trail, the 'observer effect' might have influenced students to be more performative in their critical engagement with AI than they would be in a non-monitored environment. Future research should consider longitudinal designs involving larger, multi-institutional cohorts and the integration of automated stylometry tools to further validate the 'human filter' in AI-assisted academic prose.

Conclusion

This study has demonstrated that the ethical integration of generative AI in EFL writing is not a binary choice between prohibition and total adoption, but a nuanced navigation across the Authorial Spectrum. The findings confirm that while temporal pressures often drive students toward the "Creator" model—compromising deep learning—the implementation of Socratic Prompting and Transparency Logs serves as a vital safeguard for maintaining student agency and cultural identity. By shifting the focus from the final product to a procedural ethics framework, educators can mitigate the risks of "digital ghostwriting" and "linguistic sanitization." Ultimately, this research advocates for a shift in academic integrity policies: moving away from reactive detection toward proactive Critical AI Literacy, where the preservation of the student’s unique authorial voice remains the primary measure of scholarly success.

References

Amani, M., & Abbas, N. (2024). Navigating the AI frontier: The impact of generative models on academic integrity and student voice. *Journal of Academic Ethics*, *22*(3), 315–334. https://doi.org/10.1007/s10805-024-09512-4

Bond, M., & Khosravi, H. (2025). Human-AI collaboration in higher education: A systematic review of pedagogical implications. *Computers and Education Open*, *6*(1), 100–118. https://doi.org/10.1016/j.caeo.2025.100201

Choi, J., & Lee, H. (2024). The homogenization of student prose: How LLMs influence linguistic diversity in international education. *Modern Language Journal*, *108*(2), 245–263. https://doi.org/10.1111/modl.12910

Duarte, A., & Santos, R. (2023). Digital shadows: The rise of AI-assisted ghostwriting in undergraduate submissions. *Journal of Higher Education Policy and Management*, *45*(4), 412–429. https://doi.org/10.1080/1360080X.2023.2201456

Farrelly, T., & Baker, N. (2024). Beyond detection: Re-evaluating assessment design in the age of ChatGPT. *Assessment & Evaluation in Higher Education*, *49*(6), 720–738. https://doi.org/10.1080/02602938.2024.2305612

Gao, X., & Zheng, Y. (2025). Critical AI literacy: Empowering EFL learners to maintain agency in automated environments. *Language Teaching Research*, *29*(1), 55–76. https://doi.org/10.1177/13621688251234567

Hockly, N. (2023). Artificial Intelligence in English Language Teaching: The good, the bad and the ugly. *ELT Journal*, *77*(2), 200–208. https://doi.org/10.1093/elt/ccad001

Kaur, S., & Singh, R. (2024). Ethical dilemmas of the 'Correctness' trap: AI tools and the loss of non-native speaker identity. *TESOL Journal*, *15*(3), e784. https://doi.org/10.1002/tesj.784

Liu, M., & Zhao, H. (2024). Tracking the digital footprint: Using process data to distinguish between human and AI writing. *Journal of Writing Research*, *16*(2), 145–172. https://doi.org/10.17239/jowr-2024.16.02.01

Moorhouse, B. L., & Yeo, M. A. (2023). Using AI-based writing assistants in the EFL classroom: Ethical and pedagogical considerations. *TESOL Quarterly*, *57*(4), 1320–1335. https://doi.org/10.1002/tesq.3250

Prentice, S., & Kormos, J. (2025). The psychological impact of algorithmic feedback on student writing self-efficacy. *Journal of Second Language Writing*, *67*(1), 101150. https://doi.org/10.1016/j.jslw.2025.101150

Richards, K., & Smith, L. (2024). Redefining the 'Author': Translingualism and generative AI in global academia. *International Journal of Bilingual Education and Bilingualism*, *27*(5), 580–598. https://doi.org/10.1080/13670050.2024.2210987

Sullivan, M., & Kelly, A. (2024). From correction to creation: How generative AI shifts the focus of EFL writing instruction. *System*, *120*(1), 103214. https://doi.org/10.1016/j.system.2024.103214

Suseno, E. (2024). Podcast and translation learning as a means of developing students' reading, speaking, and listening skills. *Linguistics Initiative*, *4*(2), 220–242.

Suseno, E. (2024a). The lexicogrammatical approach: Enhancing delivery skills through communicative grammar development. *JETA (Journal of English Teaching and Applied Linguistic)*, *5*(2), 47–72. https://doi.org/10.52217/jeta.v5i2.1694

Suseno, E. (2024c). Using podcasts to improve language proficiency for English as a foreign language students. *Innovare Journal of Social Sciences*, *12*(4), 1–9. https://doi.org/10.22159/ijss.2024v12i4.51673

Suseno, E. (2025a). Using the relevant person's experience and the artificial intelligence-derived materials to improve speaking during discussions. *Innovare Journal of Social Sciences*, *13*(3), 7–12. https://doi.org/10.22159/ijss.2025v13i3.53490

Suseno, E. (2025d). Using image description to influence the development of speaking abilities through artificial intelligence. *Innovare Journal of Social Sciences*, *13*(1), 18–24. https://doi.org/10.22159/ijss.2025v13i1.52857

Tan, C., & Wang, S. (2023). The transparency mandate: Students' perspectives on disclosing AI use in academic essays. *Ethics and Education*, *18*(2), 189–210. https://doi.org/10.1080/17449642.2023.2198074

Zhang, D. (2025). Prompt engineering as a new literacy: Longitudinal shifts in student critical thinking. *Educational Psychologist*, *60*(2), 88–106. https://doi.org/10.1080/00461520.2025.2405611