

From Cognitive Offloading to Reflective AI Use in EFL Writing

Trang Hong Nguyen

Master of Science in Applied Linguistics

Doctor of Education Candidate in Emerging Technologies in Education

Abstract

Generative artificial intelligence (GenAI) has become a routine source of linguistic support, feedback, and idea generation for English as a Foreign Language (EFL) learners. At the same time, its ability to complete substantial portions of academic work raises concerns about cognitive offloading and the preservation of learner judgment. This study examines actual patterns of AI use in an EFL learning context and considers how formal guidance and teacher expectations may support a shift from simple task delegation toward more reflective AI use. The study draws on a needs-assessment dataset comprising 100 learners enrolled in Basic English, Build Up, and IELTS programmes and consultation data from 35 teachers and academic staff. Descriptive statistics, cross-tabulations, and Pearson chi-square tests were used. Seventy-one percent of learners reported using AI for learning, whereas only 28% reported receiving formal guidance. Among AI users, the most common purposes were translation (76.1%), grammar explanation (70.4%), vocabulary support (63.4%), writing revision (59.2%), and idea development (53.5%); 16.9% reported using AI to complete assignments. Formal guidance was not significantly associated with overall AI use, $\chi^2(1)=3.16$, $p=.076$. Among AI users, guided learners reported more writing revision (75.0% vs. 51.1%) and less complete-assignment delegation (8.3% vs. 21.3%), although these subgroup differences did not reach statistical significance. Teacher consultation data revealed substantial inconsistency: only 25.7% reported expectations consistent with colleagues, while 57.1% reported inconsistent expectations. The findings suggest that the central educational challenge is not simply whether learners use AI, but whether institutions provide coherent guidance that preserves human reasoning, verification, and authorship in AI-supported EFL writing.

Keywords: Generative AI; EFL writing; cognitive offloading; reflective AI use; AI guidance; teacher expectations; learner agency

Date of Submission: 09-09-2026

Date of Acceptance: 18-09-2026

I. Introduction

Background of the Study

Generative artificial intelligence has rapidly altered the ecology of English-language learning. Learners can now obtain instant translation, grammar explanations, vocabulary support, idea suggestions, revision feedback, and complete text generation within a single conversational interface. In EFL writing, these affordances may reduce barriers associated with limited linguistic resources and provide timely support that would otherwise require teacher or peer assistance. Barrot [5] notes that ChatGPT and related systems can support language form, coherence, and writing practice, while Kasneci et al. [6] emphasize their potential for personalized support and interaction. However, both sources also stress accuracy, overreliance, and the need for critical evaluation.

These developments make cognitive offloading a useful theoretical lens. Cognitive offloading refers to the use of external actions or tools to reduce the internal cognitive demands of a task [4]. Offloading is not inherently harmful; learners routinely use dictionaries, notes, calculators, and digital tools to manage cognitive load. The educational concern arises when the external tool no longer supports reasoning but begins to substitute for the reasoning that the task was designed to develop. In AI-supported writing, the distinction between asking for language correction and asking an AI system to produce an entire assignment is therefore pedagogically important.

Nguyen's earlier research provides a direct foundation for this distinction. The Ethical Digital Citizenship Education framework argues that responsible participation in AI-mediated environments requires ethical awareness, responsible participation, and reflective judgment [1]. A subsequent study on emotion-sensitive AI emphasizes human autonomy, transparency, and the preservation of teacher professional judgment [2]. Most directly, Nguyen's study of 132 students found that generative-AI usage and cognitive offloading

negatively predicted critical thinking and verification practices, whereas AI literacy had a positive effect [3]. The next question is therefore practical: what kinds of guidance and classroom boundaries can help learners use AI as support without allowing it to replace the cognitive work of learning?

Statement of the Problem

The problem addressed in this study is the mismatch between widespread learner use of GenAI and the limited consistency of formal guidance. Educational debates often focus on whether AI should be permitted, but such a binary framing obscures important differences among uses. Translation, grammar explanation, vocabulary exploration, idea generation, revision, and complete-assignment generation place different demands on learner judgment and authorship. Without explicit guidance, learners may not have a clear basis for deciding which uses support learning and which uses risk excessive cognitive delegation.

The evidence pack examined in this study provides a useful opportunity to investigate this issue empirically. It records AI-use patterns from 100 learners and policy expectations from 35 teachers and academic staff. Unlike a psychometric study, the dataset does not directly measure critical thinking or cognitive offloading with multi-item scales. Accordingly, cognitive offloading is used here as a conceptual lens rather than as a measured latent variable. The empirical focus is on observable self-reported AI purposes, access to formal guidance, and the consistency of teacher expectations.

Research Questions

1. To what extent do EFL learners use AI for learning, and for which language-learning and writing-related purposes?
2. Is formal AI guidance associated with learners' overall AI use and with selected writing-related AI practices?
3. What boundaries do teachers and academic staff report for AI-supported brainstorming and language correction?
4. How consistent are teacher expectations, and what do the learner and teacher data imply for the development of reflective AI use in EFL writing?

Research Objectives

- To describe the prevalence and purposes of AI use among learners across three EFL programme types.
- To examine associations between formal guidance and selected AI-supported writing practices.
- To document teacher expectations regarding brainstorming, language correction, and consistency of AI-use boundaries.
- To interpret the findings through cognitive offloading, AI literacy, ethical digital citizenship, and human-centered AI perspectives.

Significance of the Study

The study contributes to the literature in three ways. First, it separates AI-use purposes rather than treating GenAI engagement as a single behavior. Second, it links student practice with teacher-level expectations, allowing the institutional guidance gap to be examined from both sides. Third, it extends Nguyen's earlier work [1]-[3] by moving from the identification of cognitive and ethical risks toward a practical framework for reflective use. The results may inform EFL curriculum design, teacher development, classroom AI policies, and process-based writing assessment.

Definition of Key Terms

- Cognitive offloading: the externalization of cognitive work to tools or environmental supports in order to reduce internal processing demands [4]. In this study it is a theoretical lens, not a directly measured scale.
- Reflective AI use: deliberate use of AI in which learners retain responsibility for purpose-setting, verification, revision, and final judgment.
- Formal guidance: participants' self-reported receipt of explicit instruction or guidance concerning AI use for learning.
- Writing-related AI use: AI use for idea development, writing revision, or complete-assignment completion in the supplied dataset.

II. Literature Review

Cognitive Offloading and Generative AI

Risko and Gilbert [4] conceptualize cognitive offloading as a normal strategy for changing the information-processing requirements of a task. Their framework is especially relevant to GenAI because conversational systems can externalize not only memory or calculation but also higher-order activities such as planning, explanation, argument generation, and revision. The key educational issue is therefore not the

existence of offloading but the calibration of what is offloaded. When AI handles low-level linguistic correction, the learner may preserve attention for content and organization. When AI generates the central argument, evidence, and final text, the learner may lose opportunities to practice the targeted reasoning.

Nguyen [3] empirically connected this concern to verification and critical thinking in higher education. That study found that cognitive offloading and GenAI usage negatively predicted critical-thinking and verification behaviors, while AI literacy positively predicted them. The present study does not re-measure those constructs; instead, it examines the practical usage patterns and guidance conditions that may precede more or less reflective forms of offloading.

Generative AI in EFL Writing

EFL writing is particularly susceptible to both the benefits and risks of GenAI. Barrot [5] identifies potential support for language form, coherence, and writing practice but also warns of inaccurate responses and implications for academic integrity. GenAI can act as a readily available language resource, particularly for learners who need vocabulary, grammar, or revision support. More recent work on AI-mediated scaffolding in L2 academic writing continues to emphasize the importance of integrating AI within a pedagogical design rather than treating the system as an autonomous writer [15].

Writing revision is a useful example of the distinction between assistance and substitution. AI-generated feedback can support revision frequency and writing quality when learners remain responsible for evaluating and applying suggestions [16]. By contrast, complete task generation can remove the need for learners to plan, compose, and revise. This difference provides the rationale for examining individual AI purposes instead of using a single binary indicator of AI adoption.

AI Literacy, Reflective Judgment, and Human Agency

AI literacy frameworks increasingly emphasize more than technical access. Long and Magerko [7] describe competencies required to understand and interact critically with AI, while Ng et al. [8] similarly frame AI literacy as involving knowledge, use, evaluation, and ethical understanding. UNESCO's guidance for GenAI in education calls for a human-centered approach that protects human agency and embeds ethical validation within pedagogical design [9]. Its 2024 competency frameworks for students and teachers further emphasize critical judgment, human accountability, ethics, and responsible AI pedagogy [10], [11].

These perspectives align with Nguyen's ethical digital citizenship work [1], which positions reflective judgment as a core educational objective, and with the human-centered ethical principles proposed for AI-mediated classrooms [2]; they are also consistent with broader ethical principles for responsible AI [14]. Reflective AI use in writing can therefore be understood as a practical expression of AI literacy: learners identify a purpose for AI, evaluate the output, verify important claims, reconstruct the contribution in relation to their own reasoning, and remain accountable for the final text.

Teacher Guidance and Policy Consistency

Students' decisions about AI are shaped by what teachers permit, model, and assess. The educational value of guidance depends not only on whether guidance exists but also on whether expectations are sufficiently consistent to be interpretable. Cotton, Cotton, and Shipway [12] argue that GenAI creates new academic-integrity challenges that institutions must address through assessment design and clear expectations rather than through prohibition alone. Research on EFL teachers' beliefs similarly shows a mixture of perceived usefulness and concerns about excessive reliance, academic integrity, and the displacement of established learning resources [13].

UNESCO's teacher competency framework is relevant because it explicitly treats AI pedagogy, ethics, and human-centered decision-making as professional competencies [10]. If teachers differ substantially in what they allow for brainstorming or correction, learners may receive contradictory signals about responsible use. This creates a need for shared principles that distinguish support from substitution.

Research Gap and Conceptual Positioning

Much GenAI research either measures attitudes toward AI or evaluates learning outcomes after an intervention. Less attention has been paid to the practical middle layer between these two: the distribution of everyday AI uses and the guidance environment within which learners make decisions. The present study addresses this gap through a multi-stakeholder needs-assessment design. It treats learner purposes as behavioral indicators of how AI enters the learning process and teacher responses as indicators of the policy environment. The study does not claim that any single reported purpose proves cognitive offloading. Instead, it uses the concept to interpret a continuum from targeted support to high-delegation use.

Table 1. Conceptual Mapping of Observed AI Purposes

Observed use category	Examples in dataset	Interpretive position
Linguistic support	Translation; grammar explanation; vocabulary	Primarily supports access to language resources; does not by itself indicate reduced reasoning.
Process support	Idea development; writing revision; speaking rehearsal	May support learning when learners evaluate, select, and reconstruct AI contributions.
High-delegation indicator	Complete assignments	Signals a greater transfer of task responsibility and therefore warrants closer pedagogical guidance.

III. Research Methodology

Research Design

Consistent with established research-design principles [18], the study employed a cross-sectional, multi-stakeholder needs-assessment design. The learner dataset was analyzed quantitatively to describe AI adoption, access, formal guidance, and AI-use purposes. Teacher consultation data were used to describe reported boundaries for brainstorming and language correction and to assess consistency of expectations. The design is exploratory and descriptive rather than causal.

Research Context and Participants

The supplied evidence pack contains 100 learner records collected between 12 January and 15 February 2026 and 35 teacher or academic-staff consultation records collected between 20 January and 16 February 2026. Learners were enrolled in Basic English, Build Up, and IELTS programmes. The age distribution ranged from under 18 to 35 years and older, allowing the dataset to capture both school-age and adult English learners. Teacher roles included Basic English, Build Up, IELTS, and academic/curriculum responsibilities.

Table 2. Learner Sample Characteristics (N = 100)

Characteristic	Category	n	%
Programme	Basic English	30	30.0
Programme	Build Up	30	30.0
Programme	IELTS	40	40.0
Age group	Under 18	20	20.0
Age group	18–24	45	45.0
Age group	25–34	25	25.0
Age group	35+	10	10.0
Access type	Free account	64	64.0
Access type	Premium account	18	18.0
Access type	Shared / limited access	18	18.0

Measures and Variables

The learner file includes binary Yes/No indicators for AI use, receipt of formal guidance, and seven AI-use purposes: translation, grammar explanation, vocabulary, idea development, writing revision, speaking rehearsal, and complete assignments. Programme, age group, and access type are categorical background variables. For interpretation, idea development, writing revision, and complete assignments were treated as writing-related purposes. The variable 'complete assignments' was treated as a high-delegation indicator because it represents transfer of a larger share of task completion to AI; it was not treated as a validated cognitive-offloading scale.

The teacher dataset includes categorical responses indicating whether AI-supported brainstorming is permitted (Yes/Conditional/No), whether AI language correction is permitted (Yes/Conditional/No), and whether expectations are consistent with colleagues (Yes/No/Unsure). These variables were analyzed descriptively.

Data Analysis

Analysis used frequency distributions, percentages, cross-tabulations, and Pearson chi-square tests. Chi-square tests were used to examine the association between formal guidance and overall AI use and, among AI users, the associations between guidance and selected writing-related practices. Cramér's V was reported as an effect-size index for two-by-two tables following standard statistical guidance [17]. Because the study is exploratory, p values are interpreted together with descriptive differences and effect sizes rather than as a sole decision criterion. Teacher consultation variables were summarized with counts and percentages because the teacher sample was modest and the study's primary purpose was needs assessment.

Data Integrity and Scope

The analytical workbook contains de-identified participant codes rather than names or contact information. No values were imputed or generated for the present analysis. The study is limited to variables actually contained in the evidence pack. Accordingly, it does not report psychometric scores for critical thinking, metacognition, AI literacy, or cognitive offloading, and it does not infer causal effects from cross-sectional associations.

IV. Findings and Discussion

Prevalence of AI Use and Formal Guidance

Of the 100 learners, 71 reported using AI for learning (71.0%), while 29 reported no AI use. Only 28 learners reported receiving formal guidance (28.0%). This produces a clear guidance gap: AI use was more than twice as common as formal guidance. Among the 71 AI users, 24 had received formal guidance (33.8%), meaning that approximately two thirds of active AI users reported using the technology without formal guidance.

Table 3. AI Use and Formal Guidance

Student metric	n	%
Used AI for learning: Yes	71	71.0
Used AI for learning: No	29	29.0
Received formal guidance: Yes	28	28.0
Received formal guidance: No	72	72.0

The association between formal guidance and overall AI use did not reach the conventional .05 significance threshold, $\chi^2(1, N=100)=3.16$, $p=0.076$, Cramér's $V=0.178$. The direction nevertheless shows that 85.7% of guided learners reported AI use compared with 65.3% of learners without guidance. This result should not be interpreted causally; formal guidance may be more likely to be offered to learners who are already using AI.

Purposes of AI Use

Table 4. Reported Purposes of AI Use

AI-use purpose	All learners n	All learners %	AI users n	AI users %
Translation	54	54.0	54	76.1
Grammar explanation	50	50.0	50	70.4
Vocabulary	45	45.0	45	63.4
Idea development	38	38.0	38	53.5
Writing revision	42	42.0	42	59.2
Speaking rehearsal	24	24.0	24	33.8
Complete assignments	12	12.0	12	16.9

The pattern shows that learners primarily used AI for targeted language and writing support rather than for total assignment completion. Among AI users, translation was the most common purpose (76.1%), followed by grammar explanation (70.4%), vocabulary support (63.4%), writing revision (59.2%), and idea development (53.5%). Speaking rehearsal was reported by 33.8%. Complete-assignment use was the least common purpose at 16.9% of AI users and 12% of the full sample.

Writing-related AI use was nevertheless widespread. Sixty learners used AI for at least one of the three writing-related purposes (idea development, writing revision, or complete assignments), representing 60.0% of the full sample and 84.5% of active AI users. This supports the relevance of EFL writing as a focal context: even though the evidence pack includes broader language-learning uses, most active AI users engaged with AI in the writing process.

Formal Guidance and Writing-Related AI Practices

Table 5. Formal Guidance and Selected Writing-Related Practices among AI Users

Practice	Guided AI users (n=24)	Unguided AI users (n=47)	χ^2	p	Cramér's V
Idea development	16/24 (66.7%)	22/47 (46.8%)	1.78	0.182	0.158
Writing revision	18/24 (75.0%)	24/47 (51.1%)	2.84	0.092	0.200
Complete assignments	2/24 (8.3%)	10/47 (21.3%)	1.09	0.297	0.124

The subgroup analysis suggests potentially meaningful directional differences, although none met the .05 threshold in this sample. Guided AI users reported idea development more often than unguided users (66.7% vs. 46.8%) and writing revision more often (75.0% vs. 51.1%). By contrast, guided users were less likely to

report using AI to complete assignments (8.3% vs. 21.3%). The largest of these small subgroup effects concerned writing revision (Cramér's $V=.200$, $p=.092$).

These findings should be interpreted as needs-assessment signals rather than evidence that guidance causes reflective use. The small guided subgroup limits statistical power, and guidance quality is not measured. Still, the direction is pedagogically relevant: formal guidance co-occurs with greater use of AI for revision and lower rates of complete-assignment delegation. This pattern is consistent with the idea that clearer instructional boundaries may help position AI as a support for revising and developing ideas rather than as a replacement for task ownership.

Programme Differences

AI adoption did not differ significantly across Basic English, Build Up, and IELTS programmes, $\chi^2(2, N=100)=2.34$, $p=0.310$. AI use was reported by 76.7% of Basic English learners, 76.7% of Build Up learners, and 62.5% of IELTS learners. The absence of a statistically significant programme difference suggests that GenAI use is not confined to one proficiency track and that institution-wide or programme-wide guidance is more appropriate than assuming the issue belongs only to advanced learners.

Teacher Expectations and Consistency

Table 6. Teacher Consultation Results (n = 35)

Teacher indicator	Response	n	%
Brainstorming permitted	Yes	14	40.0
Brainstorming permitted	Conditional	13	37.1
Brainstorming permitted	No	8	22.9
Language correction permitted	Yes	18	51.4
Language correction permitted	Conditional	10	28.6
Language correction permitted	No	7	20.0
Expectations consistent with colleagues	Yes	9	25.7
Expectations consistent with colleagues	No	20	57.1
Expectations consistent with colleagues	Unsure	6	17.1

Teacher responses reveal both openness to AI and substantial uncertainty about shared boundaries. Forty percent permitted AI brainstorming without qualification, while 37.1% permitted it conditionally. Language correction was treated somewhat more permissively: 51.4% responded Yes and 28.6% Conditional. The clearest institutional issue was consistency. Only 9 of 35 respondents (25.7%) reported expectations consistent with colleagues; 20 (57.1%) reported inconsistency and 6 (17.1%) were unsure.

This result helps explain why student guidance remains limited even when individual teachers are willing to use AI pedagogically. A learner may encounter one instructor who permits brainstorming, another who permits it only conditionally, and another who prohibits it. Such variability makes it difficult for students to form stable norms about authorship, acceptable assistance, and verification. The finding is closely aligned with UNESCO's call for coherent AI competencies and institutional strategies for both teachers and students [9]-[11].

Discussion: From Offloading to Reflective AI Use

Taken together, the learner and teacher results suggest that the immediate challenge is not widespread assignment outsourcing. Most AI users reported targeted support functions, and complete-assignment use was comparatively uncommon. This is important because it prevents an overly alarmist interpretation of AI adoption. The same technology that can replace cognitive work can also scaffold language access and revision. The educational task is to distinguish forms of offloading that preserve learner reasoning from forms that displace it.

The results extend Nguyen [3] in a complementary way. The earlier study demonstrated statistical relationships among GenAI usage, cognitive offloading, verification, AI literacy, and critical thinking. The present needs-assessment dataset identifies a plausible pedagogical mechanism for addressing that risk: coherent guidance. Formal guidance was limited, and teacher expectations were often inconsistent. Where guidance existed, descriptive trends favored revision and idea development over complete-assignment delegation. Although those subgroup differences were not statistically conclusive, they provide an empirical rationale for testing structured reflective-AI interventions in future research.

The findings also reinforce the ethical perspective developed in Nguyen [1] and [2]. Ethical digital citizenship requires reflective judgment rather than mere compliance, while human-centered AI requires that human agency and professional judgment remain visible. In EFL writing, these principles can be translated into concrete classroom actions: learners should think before prompting, specify why they are consulting AI, verify

important content, reconstruct AI contributions in relation to their own purpose, and remain accountable for the final text.

Proposed Reflective AI Routine for EFL Writing

Table 7. Think-Ask-Verify-Reconstruct-Reflect Routine

Stage	Learner action	Purpose
Think	Generate an initial idea, position, vocabulary set, or outline before consulting AI.	Preserve independent activation of knowledge and clarify the learning goal.
Ask	Use AI for a defined purpose such as explanation, alternatives, feedback, or revision.	Prevent vague or total task outsourcing.
Verify	Check factual claims, examples, language choices, and task fit.	Maintain epistemic responsibility and reduce uncritical acceptance.
Reconstruct	Revise, reorganize, or rewrite selected AI contributions using the learner's own reasoning and voice.	Protect authorship and integrate learning.
Reflect	State what AI contributed, what was changed or rejected, and why.	Develop metacognitive awareness and accountable AI use.

This routine is proposed as a pedagogical implication of the evidence rather than as an intervention tested by the current dataset. Its purpose is to turn general rules such as 'use AI responsibly' into observable learning behaviors. It also provides a common language through which teachers can align expectations across courses.

V. Conclusion and Recommendations

Conclusion

This study examined AI-use practices among 100 EFL learners and AI-related expectations among 35 teachers and academic staff. AI use was common (71%), but formal guidance was much less common (28%). Learners primarily used AI for translation, grammar, vocabulary, revision, and idea development, while complete-assignment generation was relatively uncommon. Among AI users, guided learners showed a descriptive tendency toward more revision-oriented use and less complete-assignment delegation, although the subgroup differences were not statistically significant. Teacher consultation data revealed a more substantial problem: only one quarter reported expectations consistent with colleagues.

The study therefore reframes the movement from cognitive offloading to reflective AI use as a guidance problem rather than a simple adoption problem. AI assistance cannot be classified as beneficial or harmful solely from frequency of use. What matters is the purpose for which AI is used, the amount of task responsibility transferred to the system, the extent to which learners evaluate its contribution, and the consistency of expectations communicated by educators. Reflective AI use offers a practical direction for protecting learner agency without denying the legitimate value of AI-supported language learning.

Educational Implications

- Institutions should develop shared AI-use language that distinguishes brainstorming, language correction, revision, and complete task generation rather than using a single permitted/prohibited rule.
- EFL writing tasks should require visible learner decisions, such as an initial outline, revision rationale, source check, or short reflection on accepted and rejected AI suggestions.
- Teacher professional development should address both technical AI literacy and the pedagogical boundary between scaffolding and substitution.
- Formal student guidance should emphasize verification, authorship, and purposeful prompting, not only plagiarism or academic misconduct.
- Programme-wide expectations are preferable to teacher-by-teacher rules because AI use was found across Basic English, Build Up, and IELTS groups.

Recommendations for Practice

1. Adopt a common Think-Ask-Verify-Reconstruct-Reflect routine for AI-supported writing.
2. Permit low-risk language support where pedagogically appropriate, while requiring stronger process evidence when AI contributes to argument, ideas, or substantive revision.
3. Ask learners to disclose the purpose of AI use rather than merely whether AI was used.
4. Require verification whenever AI generates factual, statistical, or source-related content.
5. Review teacher expectations collectively so that learners receive consistent guidance across classes and levels.

Limitations of the Study

Several limitations should be acknowledged. First, the study uses a needs-assessment dataset from a specific EFL context and does not employ probability sampling, limiting generalizability. Second, the learner measures are self-reported binary indicators rather than validated multi-item psychometric scales. Third, the study does not directly measure critical thinking, verification quality, metacognition, or cognitive offloading. Consequently, cognitive offloading is an interpretive framework, not an empirical score in this dataset. Fourth, formal guidance is recorded as a Yes/No variable without information about its content, duration, or quality. Fifth, teacher data capture reported policy expectations rather than observed classroom practice. Finally, the cross-sectional design cannot establish causal relationships.

Recommendations for Future Research

Future research should convert the present needs assessment into an intervention study. A quasi-experimental design could compare ordinary GenAI-supported writing with a structured reflective-AI routine over several weeks, using validated measures of cognitive offloading, verification behavior, metacognitive regulation, and critical-thinking performance. Process data such as prompt histories, revision traces, and independent writing samples would help distinguish improvements in final text from improvements in learning. A larger multi-site sample could also test whether shared teacher guidelines reduce complete-task delegation and increase verification-oriented AI use.

References

- [1] Nguyen, T. H. (2025). Educating for ethical digital citizenship in the AI era. *IOSR Journal of Humanities and Social Science*, 30(12, Series 2), 47-57. <https://doi.org/10.9790/0837-3012024757>
- [2] Nguyen, T. H. (2026a). Designing ethical guidelines for emotion-sensitive AI in classrooms. *IOSR Journal of Humanities and Social Science*, 31(3, Series 1), 65-78. <https://doi.org/10.9790/0837-3103016578>
- [3] Nguyen, T. H. (2026b). Assisted or replaced? Generative AI and the erosion of students' critical thinking and verification practices. *IOSR Journal of Research & Method in Education*, 16(2, Series 3), 39-57. <https://doi.org/10.9790/7388-1602033957>
- [4] Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, 20(9), 676-688. <https://doi.org/10.1016/j.tics.2016.07.002>
- [5] Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, 57, 100745. <https://doi.org/10.1016/j.asw.2023.100745>
- [6] Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., et al. (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>
- [7] Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1-16. <https://doi.org/10.1145/3313831.3376727>
- [8] Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, 100041. <https://doi.org/10.1016/j.caeai.2021.100041>
- [9] Miao, F., & Holmes, W. (2023). Guidance for generative AI in education and research. UNESCO. <https://doi.org/10.54675/EWZM9535>
- [10] Miao, F., & Cukurova, M. (2024). AI competency framework for teachers. UNESCO. <https://doi.org/10.54675/ZJTE2084>
- [11] Miao, F., Shiohira, K., & Lao, N. (2024). AI competency framework for students. UNESCO. <https://doi.org/10.54675/JKJB9835>
- [12] Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228-239. <https://doi.org/10.1080/14703297.2023.2190148>
- [13] Gao, Y., Wang, Q., & Wang, X. (2024). Exploring EFL university teachers' beliefs in integrating ChatGPT and other large language models in language education: A study in China. *Asia Pacific Journal of Education*, 44(1), 29-44. <https://doi.org/10.1080/02188791.2024.2305173>
- [14] Floridi, L., Cows, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., et al. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689-707. <https://doi.org/10.1007/s11023-018-9482-5>
- [15] Godoy, D. (2026). AI mediated scaffolding for second language academic writing in EFL composition classrooms. *Discover Education*, 5, 475. <https://doi.org/10.1007/s44217-026-01386-0>
- [16] Mekheimer, M. (2025). Generative AI-assisted feedback and EFL writing: A study on proficiency, revision frequency and writing quality. *Discover Education*, 4, 170. <https://doi.org/10.1007/s44217-025-00602-7>
- [17] Field, A. (2018). *Discovering statistics using IBM SPSS Statistics* (5th ed.). SAGE Publications.
- [18] Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.

VII. Appendix: Dataset Coding Used in the Analysis

Table A1. Variables in the Evidence Pack

Variable	Coding / categories
Used AI for learning?	Yes / No
Received formal guidance?	Yes / No
AI-use purposes	Yes / No for Translation, Grammar explanation, Vocabulary, Idea development, Writing revision, Speaking rehearsal, Complete assignments
Programme	Basic English / Build Up / IELTS
Age group	Under 18 / 18–24 / 25–34 / 35+
Access type	Free account / Premium account / Shared or limited access
Teacher brainstorming policy	Yes / Conditional / No

Teacher language-correction policy	Yes / Conditional / No
Teacher expectation consistency	Yes / No / Unsure

Note. No synthetic item-level variables were added. Statistical results in this manuscript are derived from the supplied learner and teacher records.