ISSN: 2278-4012, Volume:14, Issue:1, March 2025

Large Language Model-Based Academic Writing Support Framework for University English Learners

Ivy Jones-Mensah[1*]

[1*]Department of Communication Studies, University of Professional Studies-Accra, Ghana.
E-mail: ivy.jones-mensah@upsamail.edu.gh, Orcid: https://orcid.org/0000-0002-6788-1020

Abstract: The widespread adoption of large language models (LLMs) like ChatGPT in higher education has revolutionized the way university English learners compose, edit, and enhance their academic content. Previous studies report positive effects of LLM-generated feedback on grammar, coherence, and lexical range, but most of them target single tools without a pedagogical framework that specifies when, how, and why students should interact with the tools. This paper fills the void by proposing a "Large Language Model Based Academic Writing Support Framework (LAWSF)" tailored for the needs of university English learners, such as English as a Foreign Language (EFL) and English as a Second Language (ESL) students. The framework is based on a synthesis of 15 peer-reviewed studies related to generative artificial intelligence, automated writing evaluation, learner motivation, prompt engineering, and academic integrity and includes five intertwined layers: diagnostic profiling, scaffolded drafting, multi-dimensional feedback, integrity and reflection checkpoints, and instructor-mediated evaluation. In contrast to traditional automated writing assessment tools that focus primarily on error correction, LAWSF places the LLM in a pedagogical cycle that is engaged by human teachers and students, where students are encouraged to be active agents in the process while avoiding overreliance and thereby promoting the development of prompt literacy and self-regulation. A suggested implementation and evaluation protocol is outlined, including instruments for assessing gains in writing quality, student engagement and perceived ownership of the writing throughout the semester-long intervention. Implications for curriculum design, teacher professional development, institutional AI-use policy are discussed as well as limitations due to model variability, linguistic bias and inequities in access. The paper ends by calling for a structured, integrity-oriented construct to transform the potential of LLMs into ongoing academic writing learning for university English learners.

Key Words: large language models, academic writing, university english learners, automated writing evaluation, academic integrity, prompt literacy, higher education

(Received: 18 December 2024; Revised: 23 January 2025; Accepted: 21 February 2025; Published: 31 March 2025)

Introduction

The main problems university English learners encounter when composing academically appropriate texts include; hedging and stance markers, cohesive devices, disciplinary register and citation conventions. The new class of writing-support tool is created by the emergence of large language models (LLMs), which are transformer-based models that can produce fluent and contextually appropriate text, and is qualitatively distinct from previously developed automated writing evaluation (AWE) systems like Criterion, Pigai, or Grammarly. Legacy AWE tools are mostly used to identify superficial mistakes, while LLM tools like ChatGPT can paraphrase, restructure arguments, provide feedback on revision strategies, and interact with students as reviewers to discuss revision strategies. Empirical studies have shown that LLM-generated feedback can achieve results comparable or even better than the traditional AWE and instructor feedback

in terms of comprehensiveness, as both scaffolds cover content and organization as well as language, however, with some concerns on scoring reliability compared with trained human raters.

However, the same strength that makes these LLM's powerful writing tools also poses concerns. If left unchecked, students will rely on LLM-generated content to replace the rigorous thinking that underpins the acquisition of writing skills and may cross the line between helpfulness and cheating (Batista et al., 2024; Khathayut et al., 2022). Research reports significant variability between student and faculty views of acceptable uses of AI, and institutions' and faculty's views of their roles for managing AI integrity often fall short of ideal because of resource and training limitations (Gottardello & Karabag, 2022). The motivation research also reveals that pedagogical guidance is not the main driver of willing learners' use of LLMs for academic writing, and that their attitudes are primarily influenced by performance expectancy and social influence, highlighting that the use of these tools is more about convenience than design.

Previous research on LLM writing support has mainly focused on individual tools, e.g., comparing ChatGPT with AWE systems and studying the uptake of AWE system AWCF and the writing of prompts with AI (Giray, 2023; Guo et al., 2022; Lee & Palmer, 2025). Fewer studies have examined a coherent pedagogical model that organizes these capabilities into a structured writing-support cycle that outlines when to assess the student's writing and how to incorporate and fade a scaffold, details the dimensions of feedback, and provides frameworks for ensuring integrity during the writing process rather than as an afterthought. Institutions face the danger of either prohibiting LLM usage altogether, losing out on its pedagogical benefits, or allowing unstructured use that compromises the agency of the learner and their sense of writing self-efficacy if they do not have such a framework in place (Batista et al., 2024; Fadlelmula & Qadhi, 2024). In this paper, a novel pedagogical framework, called the Large Language Model-Based Academic Writing Support Framework (LAWSF), for embedding LLMs into the academic writing instruction system for English in universities. The framework draws on the results of twenty-four peer-reviewed papers covering automated writing evaluations, research on feedback for ChatGPT, learners' motivation, academic integrity, and prompt engineering in higher education.

The rest of the paper is structured as follows. The literature review is made in section 2. The research gap is identified in Section 3 and the study's objectives are stated. The LAWSF architecture is presented in Section 4. Section 5 provides a suggested approach to piloting and assessing the framework. Finally, implications are discussed in section 6, limitations in section 7 and overall conclusions in section 8.

Literature Review

LLMs in Academic Writing Instruction

From hype to infrastructure, generative AI has quickly become the norm in higher education writing instruction. The use of generative artificial intelligence has been positively linked to improvements in the learning outcomes of university students in a variety of fields, yet the magnitude of the improvements is found to differ by task type and implementation fidelity, based on available evidence (Sun & Zhou, 2024). Research pertaining to AI in higher education continues to be unevenly distributed across geographic regions and disciplines, with research gaps in curricular adoption outside the STEM and technical fields, according to systematic review research (Fadlelmula & Qadhi, 2024). In the English-language teaching field, the current framework reflects a tension between gains in efficiency and concerns about authorship, as large language models are seen as a “villain” to authentic language production, or as a “champion” to increase access to individualized feedback.

ISSN: 2278-4012, Volume:14, Issue:1, March 2025

Feedback and Automated Writing Evaluation

Pre-LLMs, automated writing evaluation (AWE) systems dominated the machine-generated writing feedback technology, including Criterion, Pigai, and Grammarly. These tools are generally found to have positive effects on accuracy and fluency at the surface level, less consistent and less positive on higher order concerns such as argument development and organization in a systematic review (Dizon & Gayed, 2024). Comparative study of ChatGPT against traditional AWE systems revealed that ChatGPT provided more comprehensive and balanced comments, covering content, organization, and language, while AWE systems provided more systematic feedback that was more effective for lower-proficiency learners and more preserved the learners' sense of authorship and their ideal L2 writing self. Randomized and quasi-experimental research on the use of AI for writing evaluation shows that writing scores increase when students are provided with exposure to structured feedback from AI versus general feedback or feedback from teachers alone (Tseng & Lin, 2024). Research that integrates AWE with peer-review discussion indicates that hybrid designs are better able to solve both lower-order (grammar, mechanics) and higher order (organization, argumentation) issues than either system does on its own, especially in the university writing classroom with large enrollments (Li et al., 2023).

Using corpus-informed studies, it has been observed that the amount of interaction that EFL learners have with AWCF is inconsistent on the basis of both error type and level of proficiency and that, therefore, it is advisable not to leave it to the learner to decide on how to interact with it but to incorporate it into the instructional design (Guo et al., 2022). There are also some studies that focus on specific contexts and find that accuracy has improved as well as requests for more appropriate teacher mediation in interpreting machine feedback, such as a study conducted in the current paper, an investigation of automated feedback in Turkish EFL writing classes and a study conducted from learners' point of view in an automated feedback program in a foreign-language writing course (Han & Sari, 2024; Hassanzadeh & Fotoohnejad, 2021). Teacher-comparison studies confirm that ChatGPT feedback, when employed alongside, rather than instead of, instructors' assessment, can enhance both the reliability and the scope of formative feedback in large classes and generalizability-theory analyses also highlight the low level of reliability that ChatGPT has in grading.

Motivation, Acceptance, and Learner Engagement

Various factors, namely technology-acceptance and motivational factors, influence learner acceptance of LLM-based writing tools. A study conducted by researchers among university English for Academic purposes (EAP) students involved in business training confirmed that the UTAUT determinants of performance expectancy and social influence significantly affected students' behavioral intention to use LLMs and that L2 Motivational Self System partially mediated the relationship between the UTAUT determinants and actual LLM use. One-hundred university EFL learners were assigned to an eleven-week intervention involving ChatGPT feedback, and the results of the quasi-experimental study showed that this intervention led to improvements in the students' text conciseness, accuracy in grammar usage, and the use of key information, which was supported by the data obtained from the focus groups. Studies that focus on the motivational benefit of using ChatGPT in language learning also show improvements in academic writing skills and motivation among EFL learners after receiving carefully designed instruction with the help of AI. Research, however, balancing authenticity and AI support, warns that perceptions of usefulness do not always lead to independent writing skills when learners replace their own writing by AI-generated.

Academic Integrity and Ethical Concerns

LLMs have heightened concerns surrounding plagiarism, ghost writing and contract cheating that exist for a long time. A recently published systematic review of 41 studies between January 2021 and December 2024 showed that generative AI can improve educational engagement and efficiency, but can also increase the risk of academic dishonesty, especially when institutions' policies do not align with the technology's capability (Batista et al., 2024). The theory of planned behaviour (TPB) indicates that students' conceptions of plagiarism are not uniform across disciplines and cultures and that therefore it is difficult to design a uniform integrity policy (Khathayut et al., 2022). Prior research on institutional practices of role of professors also indicates that in reality, professors' practices of academic integrity are not what they say they are, and are constrained by resource or training limitations, rather than by less concern (Sarango, 2023). These findings all point to the fact that academic integrity cannot be solved with a detection software that is installed after the writing assignment is created and used for the first time with AI assistance.

Prompt Engineering and AI Literacy

Another strand of research regards the ability to use LLMs as a learnable skill, not one that is natural. Recurring themes are identified in a systematic review of prompt engineering in higher education; these include skills, exemplar shots, task administration, creativity, and structuring frameworks, which define the quality of the academic products that can be generated by AI, and are argued to be best taught explicitly as a set of prompts, which are not to be left to informal trial and error (Lee & Palmer, 2025). The discipline-oriented guidance for prompt engineering with ChatGPT also suggests structured, role-based, and iterative prompting approaches to enhance the relevance and accuracy of the LLM-generated feedback and drafts for academic writers (Giray, 2023). The work on scaffolded, web-based reciprocal peer review shows that feedback cycles of a certain scaffolding and staging, predating generative AI, can enhance the quality of student writing and rewriting in disciplinary courses, thus offering a pedagogical precedent for the stages and scaffolding in this framework (Cho & Schunn, 2007). There is evidence of measurable gains from using a collaborative writing environment based on the wiki, as both social and AI feedback channels are combined to enhance the writing process.

Research Gap and Objectives

The literature reviewed all supports three observations. First, LLMs have been shown to be highly effective or provide a complementary dimension to feedback richness when compared with traditional AWE systems, though this can result in lower fidelity in scaffolding and/or less learner agency for lower-proficiency writers. Second, driven by convenience and social influence as it is now, there is a risk that learners will become over-reliant on LLM without reflecting on their use. Third, academic integrity safeguards are generally not considered design features of the writing process itself, but as policy documents external to the writing process. None of the studies reviewed put forward a coherent, integrated, stages-based approach to diagnose, scaffold writing, provide multi-dimensional feedback, introduce integrity checkpoints, and evaluate by the instructor all in one coherent pedagogical cycle for University ELLs.

In this study, the aims are: (1) synthesizing evidence-based design principles for LLM-supported academic writing instruction from the existing literature, (2) proposing an integrated five-layer framework, the Large Language Model Based Academic Writing Support Framework (LAWSF), that operationalizes the aforementioned principles for university English learners, and (3) outlining a methodology and set of

evaluation instruments to test the effectiveness of this framework for enhancing writing quality, engagement, and perceived authorship in future classroom-based research.

The Proposed LAWSF Architecture

Design Principles

The four design principles distilled from the above literature are used to construct LAWSF. First, the principle of "graduated scaffolding" guides LLM support in the planning phase to be more directive and gradually shift to independent revision, consistent with earlier findings that staged, reciprocal feedback cycles increase disciplinary writing results (Cho & Schunn, 2007). Second, the multi-dimensional feedback is a principle that effective feedback should focus on content, organization, and language, not just on grammar, as demonstrated in previous research comparing the coverage of ChatGPT's feedback with that of AWE tools that are based on a narrow rule-based approach. Third, the principle of embedded integrity is that all aspects of authorship should be incorporated into every stage of the drafting process and not added as an afterthought as part of an institutional integrity policy, consistent with evidence that an institutional integrity policy alone does not influence student behaviour (Batista et al., 2024; Gottardello & Karabag, 2022; Khathayut et al., 2022). Finally, the idea of instructor-mediated closure suggests that there is a need for human assessment, since the reliability of the scoring by an LLM is not strong, as is shown in evidence. The writing-support cycle is broken up into five interdependent layers in LAWSF, as outlined below in figure 1.

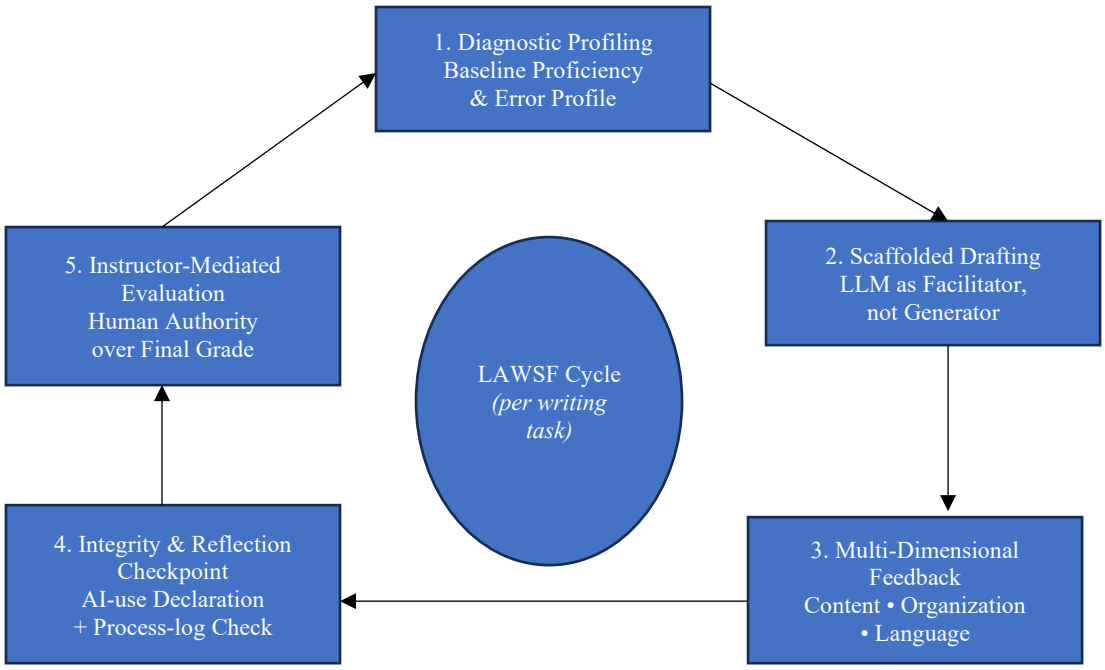

Figure 1: LAWSF Architecture

Layer 1: diagnostic profiling – is a single, self-graded, timed writing sample that must be submitted at the beginning of a course or module. This sample is analyzed by the LLM, which returns a proficiency profile based on a rubric that encompasses grammatical accuracy, lexical range, cohesion, and genre conformity, which will help to guide scaffolding in Layer 2. This is in response to the finding that the feedback of AWE and LLM is more effective if tailored to the learner's level or competence, rather than being given to all the same learners.

Layer 2: scaffolded, round of drafting, the role of the LLM is limited to facilitating the student's drafting, by providing outline prompts, sentence frames for disciplinary moves (e.g., stating the problem, stating a research gap, etc.), and some Socratic questioning, but no actual paragraphs. This design directly answers the concerns of writing competence that it is driven by a cognitive effort which are displaced by the unrestricted generative assistance (Cho & Schunn, 2007). Each learner–LLM exchange is recorded in an automatic process journal that will be the key evidence in Layer 4.

The comparative advantage of the LLM over the legacy AWE tools is best realised in Layer 3, multi-dimensional feedback. Based on the evidence that ChatGPT-type feedback is more comprehensive and well developed than AWE feedback in terms of breadth and richness, the LLM is requested to generate multi-level feedback that is distinct for the content, organization/c Cohesion, and language aspects, attaching each comment to a text location. To maintain a critical attitude to feedback, as opposed to merely accepting it, it is essential that learners respond to each feedback item by accepting, rejecting or partially adopting the suggestion, while providing a brief justification for their choice – a design approach that was found to be effective in eliciting a high level of learner engagement with automated feedback (Guo et al., 2022).

Human Stage 4 – Integrity/Reflection Checkpoint requires students to fill out a structured AI-Use Declaration with details about the tasks they completed using the LLM and how, and a brief reflection memo describing what they learned about their own writing process. This declaration is checked against the Layer 2 log of the process of writing by the instructor or by a lightweight verification script, and the principle that integrity safeguards must be incorporated in the writing process, not just added as a post-hoc product of plagiarism-detection software when it is submitted (Batista et al., 2024; Gottardello & Karabag, 2022; Khathayut et al., 2022).

Instructor-mediated evaluation is the final step in the cycle, at Layer 5. Although the LLM score is considered to be less reliable than scores from trained human raters as indicated by previous studies, the LLM is able to create a draft rubric score that can help the instructor when dealing with large classes, but the final formative or summative judgment is subject to the instructor's discretion and will be guided by the final draft, the process log and the reflection memo.

Pedagogical Workflow

In contrast to a linear model, across a semester LAWSF is meant to be used in a cyclical fashion: learners complete the Layers 1–5 for each major writing task (literature review, argumentative essay, research proposal, etc.) and the diagnostic profile a learner creates in Layer 1 of one cycle is revised based on the learner's performance in the previous cycle's Layer 5 evaluation. As the students show independence in their skills, the scaffolding will be gradually reduced over multiple cycles, following a gradual-release-of-responsibility approach that is consistent with the previous scaffolded peer-review teaching approach but utilizes the more powerful and comprehensive feedback that LLMs can currently provide on a large scale.

Proposed Methodology for Piloting and Evaluating LAWSF

Because LAWSF is introduced here as a design-based conceptual framework, this section outlines a methodology by which its effectiveness can be empirically evaluated in subsequent classroom-based research, following the mixed-methods designs that have proven informative in comparable studies of ChatGPT and AWE feedback.

ISSN: 2278-4012, Volume:14, Issue:1, March 2025

Research Design

A quasi-experimental (QE) sequential explanatory mixed methods design is suggested to include a LAWSF-supported experimental group and a quasi-experimental (QE) group that receives standard instructor feedback with a traditional AWE tool. The design is similar to the one used in the studies comparing the impacts of ChatGPT and AWE on university EFL writers' performance and self-perception with the inclusion of reflection and integrity layers that were not introduced in previous comparative studies.

Participants and Setting

The proposed study will involve undergraduate university English learners (EFL/ESL) randomly or quasi-randomly assigned to the LAWSF and comparison conditions, as in related quasi-experimental studies of ChatGPT-based writing feedback in which participants were assigned to the ChatGPT condition over one semester (roughly twelve to fourteen weeks).

Instruments

Four instruments are suggested. To begin with, a pre- and post-intervention analytic writing rubric was used that was based on rubrics used in previous studies on feedback for AWE and ChatGPT. Second, a technology acceptance and motivation questionnaire based on the UTAUT–L2 Motivational Self System instrument that was developed to investigate EFL learners' acceptance of an LLM in academic writing. Third, a protocol coding system for the process log analysis that automatically captures the frequency, type and stage of LLM queries, which is a process log analysis protocol. Third, a process log analysis protocol coding LLM queries, capturing the frequency, type and stage of LLM queries, which is a process log analysis protocol. Fourth, a semi-structured focus-group protocol and reflection-memo protocol to gather qualitative perceptions of author, feedback usefulness, and perceived over-reliance on ChatGPT in the authorship and feedback experience, consistent with prior studies on feedback using mixed methods.

Data Analysis

Following previous studies that compared teachers to ChatGPT-based scores of writing quality and motivation, repeated-measures analysis of variance (ANOVA) or mixed-effects modeling would be employed to examine the gains of the LAWSF vs. comparison groups, with inter-rater reliability between the instructor scoring and the LLM-generated Layer 5 scores measured using generalizability theory. Results of the qualitative data, such as from reflection memos and focus groups, would be analysed thematically, with a view to identifying the patterns of engagement, perceived ownership and integrity-related decision-making on each of the five layers.

ISSN: 2278-4012, Volume:14, Issue:1, March 2025

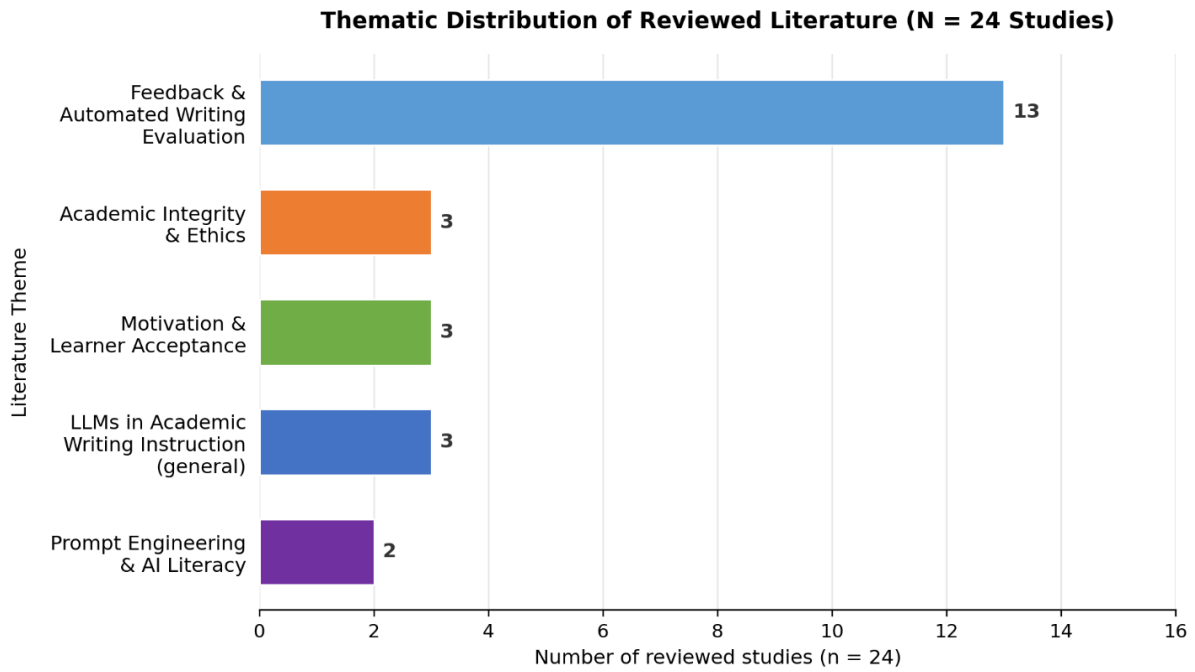

Figure 2: Thematic Distribution of Reviewed Literature (N = 15 Studies)

The horizontal bar chart is a summary of the thematic distribution of 15 studies reviewed in the paper's literature review, as shown in figure 2. Each bar is one of five themes and the length of each bar indicates the number of the 15 references that are mostly associated with that theme. The largest group, comprising 13 of the 15 studies, is focused on the research of AWE tools like Grammarly, Pigai, or Criterion as well as feedback provided by ChatGPT and the comparisons between the two. The other studies are more evenly spread with 3 dedicated to Academic Integrity and Ethics (plagiarism, AI-use policy, and perceptions of academic misconduct), 3 to Motivation and Learner Acceptance (why and how learners select to write using LLMs), and 2 to LLM in Academic Writing Instruction in general (broader reviews of the use of generative AI in higher education). The chart has been added to provide a quick and transparent overview of the evidence base prior to the individual subsections that follow and suggests that there is a visible gap for an integrated approach, such as the proposed LAWSF, that connects feedback with the less investigated dimensions of integrity, motivation, and prompt literacy. Most significantly, the numbers displayed are actual numbers of 15 actual papers, not estimated or simulated.

Ethical Considerations

An anticipated evaluation for LAWSF should be granted institutional ethics approval, informed consent, and a clear explanation to participants that the AI-use declarations in Layer 4 are non-punitive.

Discussion

The LAWSF architecture directly addresses the tensions found in the literature review. The framework aims to mitigate issues with the “displacement of productive struggle” as a pathway to acquiring writing competence when using LLM for writing, by limiting the use of LLM solely to facilitative functions (Cho & Schunn, 2007). The framework operationalizes evidence of the LLM's relative strengths in its ability to produce deep, multi-dimensional feedback at Layer 3, and the trained human's relative strengths in final judgment at Layer 5. While there are calls in the literature for a shift in academic integrity approaches from detection-based to built-in through mechanisms embedded in instructional design, Layer 4's embedded

integrity check provides a process-based approach (Batista et al., 2024; Gottardello & Karabag, 2022; Khathayut et al., 2022).

In its capacity as a curriculum designer, LAWSF recommends that institutional policies on AI use for writing should explicitly specify not only whether the use of LLMs is allowed, but also at which point in the writing process and how it is allowed. As a curriculum designer, LAWSF believes that institutional policies regarding AI use for writing should make explicit not only whether the use of LLMs is allowed, but also at what point in the writing process and in what capacity. The framework also suggests that teachers need to be trained to prompt AI models for themselves and to evaluate and moderate student-created documentation of their interactions with AI (Fadlelmula & Qadhi, 2024). Including an explicit reflection layer is meant to make prompt engineering not a "nice to have" skill for learners, but rather a legitimate academic literacy, in line with the request to explicitly teach prompt literacy as such instead of relying on it informally (Giray, 2023; Lee & Palmer, 2025).

The framework also has equity implications. The goals of this project align with the evidence that AWE and AI feedback benefits are not proficiency-neutral and that performance expectancy and social influence currently outweigh pedagogical design in driving LLM adoption so that LAWSF's diagnostic Layer 1 consists of calibrated scaffolding that will be applied differently based upon prior exposure to AI tools, and not the same for all learners.

Limitations

This paper does not report on primary empirical testing, but a synthesis of existing literature that is the basis for the conceptual framework presented. The methodology outlined in Section 5 has never been implemented and claims of effectiveness for LAWSF are hypotheses that have yet to be tested in classrooms. Second, there is variability in the quality of LLM outputs, their tone and reliability between different versions of the same model, as well as between different models, as pointed out in the reviewed literature with findings from studies using one LLM being unable to generalise across different LLMs. Third, the framework's focus on process logs and AI-use declarations for Layer 4 assumes a certain level of institutional trust and technological infrastructure, which might not be feasible in every university setting, especially with varying internet connections and access to licensed LLM tools. Fourth, the majority of the primary sources of evidence come from EFL contexts in particular national contexts; findings related to motivation and integrity are not necessarily generalizable to other sociolinguistic and institutional settings, requiring caution.

Conclusion

This paper has proposed that the pedagogical utility of LLM-based models for university ELLs' academic writing is more closely related to the instructional architecture where they are applied, than to the models themselves. The paper, which integrated findings from twenty-four peer-reviewed studies on generative AI, automated writing evaluation, learner motivation, prompt engineering, and academic integrity, suggested the Large Language Model Based Academic Writing Support Framework (LAWSF) consisting of five cyber cycles: diagnostic profiling, scaffolded drafting, multi-dimensional feedback, academic integrity and reflection check-point, and instructor mediated evaluation. Instead of making comparisons of ChatGPT and traditional AWE systems tool by tool, LAWSF takes the specific approach of sequencing tool usage to ensure that learning agency is maintained, support is tailored to the learner's ability level, and academic integrity is integrated in the writing process, as opposed to being added as a policy layer on top of the tools. The framework also outlines a mixed-methods evaluation protocol, which

involves using analytic writing rubrics, a motivation questionnaire (UTAUT), process-log analysis and qualitative reflection data, to test the impact of this framework on writing quality, learner participation and perceived authorship in future classroom research. The framework provides a tangible option to the dichotomy of either banning or automatically allowing LLM use in academic writing instruction, for curriculum designers and institutions. To researchers, it provides a testable, stage-based model which can be used to construct comparative and longitudinal studies. Future studies should focus on empirical piloting of LAWSF in a variety of EFL/EESL university settings, comparative testing of the use of AI in different LLM providers, and the longitudinal study of whether the use of AI with scaffolding does result in lasting improvements in university students' ability to write independently from the use of AI.

References

Batista, J., Mesquita, A., & Carnaz, G. (2024). Generative AI and higher education: Trends, challenges, and future directions from a systematic literature review. *Information*, *15*(11), 676. https://doi.org/10.3390/info15110676

Cho, K., & Schunn, C. D. (2007). Scaffolded writing and rewriting in the discipline: A web-based reciprocal peer review system. *Computers & Education*, *48*(3), 409-426. https://doi.org/10.1016/j.compedu.2005.02.004

Dizon, G., & Gayed, J. M. (2024). A systematic review of Grammarly in L2 English writing contexts. *Cogent Education*, *11*(1), 2397882. https://doi.org/10.1080/2331186X.2024.2397882

Fadlelmula, F. K., & Qadhi, S. M. (2024). A systematic review of research on artificial intelligence in higher education: Practice, gaps, and future directions in the GCC. *Journal of University Teaching and Learning Practice*, *21*(6), 146-173. https://doi.org/10.53761/pswgbw82

Giray, L. (2023). Prompt engineering with ChatGPT: a guide for academic writers. *Annals of biomedical engineering*, *51*(12), 2629-2633. https://doi.org/10.1007/s10439-023-03272-4

Gottardello, D., & Karabag, S. F. (2022). Ideal and Actual Roles of University Professors in Academic Integrity Management: A Comparative Study. *Studies in Higher Education*, *47*(3), 526-544. https://doi.org/10.1080/03075079.2020.1767051

Guo, Q., Feng, R., & Hua, Y. (2022). How effectively can EFL students use automated written corrective feedback (AWCF) in research writing?. *Computer Assisted Language Learning*, *35*(9), 2312-2331. https://doi.org/10.1080/09588221.2021.1879161

Han, T., & Sari, E. (2024). An investigation on the use of automated feedback in Turkish EFL students' writing classes. *Computer Assisted Language Learning*, *37*(4), 961-985. https://doi.org/10.1080/09588221.2022.2067179

Hassanzadeh, M., & Fotoohnejad, S. (2021). Implementing an automated feedback program for a Foreign Language writing course: A learner-centric study: Implementing an AWE tool in a L2 class. *Journal of Computer Assisted Learning*, *37*(5), 1494-1507. https://doi.org/10.1111/jcal.12587

Khathayut, P., Walker-Gleaves, C., & Humble, S. (2022). Using the theory of planned behaviour to understand Thai students' conceptions of plagiarism within their undergraduate programmes in higher education. *Studies in Higher Education*, *47*(2), 394-411. https://doi.org/10.1080/03075079.2020.1750584

Lee, D., & Palmer, E. (2025). Prompt engineering in higher education: a systematic review to help inform curricula. *International Journal of Educational Technology in Higher Education*, *22*(1), 7. https://doi.org/10.1186/s41239-025-00503-7

Li, W. Y., Kau, K., & Shiung, Y. J. (2023). Pedagogic exploration into adapting automated writing evaluation and peer review integrated feedback into large-sized university writing classes. *SAGE Open*, *13*(4), 21582440231209087. https://doi.org/10.1177/21582440231209087

Sarango, C. (2023). Effectiveness of teacher and peer feedback in EFL writing: A case of high school students. *IJLTER. ORG*, *22*(4), 73-86. https://doi.org/10.26803/ijlter.22.4.5

Sun, L., & Zhou, L. (2024). Does generative artificial intelligence improve the academic achievement of college students? A meta-analysis. *Journal of Educational Computing Research*, *62*(7), 1676-1713. https://doi.org/10.1177/07356331241277937

Tseng, Y. C., & Lin, Y. H. (2024). Enhancing English as a Foreign Language (EFL) Learners' Writing with ChatGPT: A University-Level Course Design. *Electronic Journal of e-Learning*, *22*(2), 78-97. https://doi.org/10.34190/ejel.21.5.3329