

## Enhancing English Language Learning through Artificial Intelligence: A Practical Framework for Improving Writing Accuracy, Feedback Quality, and Learner Autonomy

Aisha Mohamed Ahmed \*

European Academy of science and development, Turkey

\*Email (for reference researcher): [aisha.m.ahmed@gmail.com](mailto:aisha.m.ahmed@gmail.com)

تعزيز تعلم اللغة الإنجليزية من خلال الذكاء الاصطناعي: إطار عمل عملي لتحسين دقة الكتابة وجودة التغذية الراجعة واستقلالية المتعلم

عائشة محمد أحمد \*

الأكاديمية الأوروبية للعلوم والتنمية، تركيا

Received: 11-04-2025; Accepted: 23-09-2025; Published: 12-10-2025

### Abstract:

Writing in English is challenging for many adult learners of English as a Foreign Language (EFL). Errors in grammar and style persist, and traditional teacher feedback often comes with delays. Artificial intelligence (AI) offers new possibilities to enhance writing instruction by providing immediate, personalized feedback. This paper presents a practical framework for integrating AI into EFL writing pedagogy to improve writing accuracy, feedback quality, and learner autonomy. The framework combines AI-driven tools for instant corrective feedback with guided revision and teacher support. A mixed-method evaluation was conducted with adult EFL learners globally to assess the framework's effectiveness. Quantitative results show that learners who received AI-assisted feedback achieved higher grammatical accuracy and made more revisions to their writing than those with only teacher or no feedback. The AI-supported group's writing improved not only in error reduction but also in content organization and coherence, while their autonomy scores (self-directed learning measures) increased significantly. Qualitative feedback from learners indicates reduced anxiety, greater confidence, and a preference for a blended feedback approach. These findings suggest that carefully deployed AI tools can act as effective "digital writing coaches," improving the precision of writing and empowering learners to take more control of their learning. The paper discusses pedagogical implications, such as the need for teacher moderation and strategies to prevent overreliance on AI.

**Keywords:** Artificial Intelligence, EFL Writing, Writing Accuracy, Feedback Quality, Learner Autonomy, Automated Feedback.

### المخلص:

تشكل الكتابة باللغة الإنجليزية تحديًا للعديد من متعلمي اللغة الإنجليزية كلغة أجنبية (EFL) البالغين. لا تزال الأخطاء النحوية والأسلوبية قائمة، وغالبًا ما تُصاحبها ملاحظات المعلم التقليدية تأخيرات. يُنتج الذكاء الاصطناعي إمكانيات جديدة لتحسين تعليم الكتابة من خلال توفير ملاحظات فورية ومخصصة. تُقدم هذه الورقة إطارًا عمليًا لدمج الذكاء الاصطناعي في منهجية الكتابة في اللغة الإنجليزية كلغة أجنبية لتحسين دقة الكتابة وجودة الملاحظات واستقلالية المتعلم. يجمع الإطار بين أدوات مُعتمدة على الذكاء الاصطناعي لتقديم ملاحظات تصحيحية فورية مع مراجعة مُوجهة ودعم المعلم. أُجري تقييم مُختلط الأساليب على متعلمي اللغة الإنجليزية كلغة أجنبية البالغين عالميًا لتقييم فعالية الإطار. تُظهر النتائج الكمية أن المتعلمين الذين تلقوا ملاحظات بمساعدة الذكاء الاصطناعي حققوا دقة نحوية أعلى وأجروا مراجعات أكثر على كتاباتهم مقارنةً بمن تلقوا

ملاحظات المعلم فقط أو بدونها. تحسنت كتابة المجموعة التي تلقت ملاحظات الذكاء الاصطناعي ليس فقط في تقليل الأخطاء، ولكن أيضًا في تنظيم المحتوى وتماسكه، بينما ارتفعت درجات الاستقلالية (مقاييس التعلم الذاتي) لديهم بشكل ملحوظ. تشير التغذية الراجعة النوعية من المتعلمين إلى انخفاض القلق، وزيادة الثقة، وتفضيلهم لنهج التغذية الراجعة المختلط. تشير هذه النتائج إلى أن أدوات الذكاء الاصطناعي المستخدمة بعناية يمكن أن تكون بمثابة "مدربين رقميين للكتابة" فعالين، مما يحسن دقة الكتابة ويؤكد المتعلمين من التحكم بشكل أكبر في تعلمهم. تناقش الورقة البحثية الآثار التربوية، مثل الحاجة إلى إشراف المعلم واستراتيجيات منع الاعتماد المفرط على الذكاء الاصطناعي.

**الكلمات المفتاحية:** الذكاء الاصطناعي، كتابة اللغة الإنجليزية كلغة أجنبية، دقة الكتابة، جودة التغذية الراجعة، استقلالية المتعلم، التغذية الراجعة الآلية.

## Introduction

Adult learners around the world strive to improve their English writing skills for academic and professional success. Writing in a second language is difficult because it requires accuracy in grammar and vocabulary as well as the ability to organize ideas coherently. In traditional classrooms, teachers provide feedback on student writing, pointing out errors and offering suggestions. This feedback is crucial for learning, but it is often delayed and limited by class sizes and teacher workload. Many EFL learners get feedback days or weeks after submitting their work, by which time the impact on their immediate writing process is reduced. In contrast, modern artificial intelligence (AI) tools can analyze text and deliver instant feedback on language errors and style. This immediacy is a clear advantage of AI-assisted feedback over the slower pace of human feedback (Mekheimer, 2025). AI-based writing assistants (for example, grammar checkers or chatbot tutors) are becoming widely accessible, and they offer individualized suggestions on grammar, spelling, and style in real time. These tools focus primarily on surface-level correctness, helping learners notice and correct errors in their writing immediately. The potential benefits of such tools have generated considerable interest in language education (Ali, 2020; Crompton et al., 2024) (Mekheimer, 2025).

Despite these advantages, important questions remain about how AI can be used to truly enhance writing accuracy (error-free writing), feedback quality (the usefulness and depth of comments learners receive), and learner autonomy (students' ability to self-direct their learning). Educators are cautious because AI feedback tends to focus on grammar and mechanics, possibly neglecting higher-order writing skills. Teacher feedback, on the other hand, often addresses content development, argumentation, and organization – areas where human insight is valuable. Moreover, sole reliance on AI might lead to overreliance, where learners accept suggestions uncritically and do not develop their own editing skills (Mekheimer, 2025). The key is finding a balance that combines the strengths of AI and human feedback.

In this paper, we propose a practical framework for integrating AI into EFL writing instruction that aims to improve writing accuracy while also enhancing the quality of feedback and fostering greater learner autonomy. The framework is designed for adult EFL learners in diverse contexts (global settings, including under-resourced environments). It outlines how teachers and learners can collaborate with AI tools to create a supportive writing feedback cycle. The framework leverages AI's instant corrective feedback and the teacher's guidance on content and critical thinking. We also test this approach through an empirical study involving adult EFL writers to evaluate outcomes in accuracy, feedback effectiveness, and autonomy. The results provide evidence that AI, when used thoughtfully, can significantly improve learners' grammatical accuracy and independence (Mansoor et al., 2025). Students using AI feedback made more frequent revisions and reported feeling more confident and less frustrated in the writing process (Mekheimer, 2025). At the same time, the role of the teacher remains crucial to address areas beyond the scope of AI (such as idea development) and to ensure feedback is high-quality and pedagogically sound.

The remainder of this paper is organized as follows: first, we review relevant literature on AI in language learning, automated feedback, and learner autonomy. Next, we detail the proposed framework and methodology of our study. We then present and discuss the results, including

improvements in writing accuracy, feedback quality, and autonomy observed in the AI-assisted learning environment. Finally, we conclude with implications for educators and suggestions for future research on integrating AI into language education.

## Literature Review

### AI in EFL Writing Instruction and Feedback

The use of AI in language learning has evolved rapidly. Early forms of technology-assisted writing feedback, often called Automated Writing Evaluation (AWE) systems, have been in use for decades (e.g., Criterion or Grammarly). These systems automatically identify grammatical errors and sometimes provide holistic scores on writing. Modern AI tools, powered by machine learning and natural language processing, have greatly improved the accuracy and scope of automated feedback. They can detect subtle grammar mistakes, suggest rephrasing for clarity, and even generate sample improvements. For example, an AI writing assistant can underline a verb tense error and immediately suggest the correct tense. Research has shown that such immediate feedback can raise learners' awareness of errors (Dodigovic, 2005) and reduce the time between writing and correction, which is critical for effective learning (Chen et al., 2024). According to Mekheimer (2025), these tools provide immediate, individualized feedback on surface-level aspects of writing like grammar and spelling. This instant feedback contrasts with the potentially delayed feedback from teachers, which might come days later due to workload constraints. Immediate corrective feedback helps learners apply suggestions right away in their drafts, reinforcing learning at the point of need.

However, AI-generated feedback has limitations. It is typically strongest in addressing language accuracy and style. Much of the automated feedback is on *form* (e.g. correcting a tense or article usage) rather than *content*. A common critique is that AI tools “excel at identifying grammatical errors and stylistic inconsistencies” but are less effective in giving feedback on “higher-order writing aspects, such as content development and coherence” (Mekheimer, 2025). In contrast, human teachers provide feedback that spans from lower-order concerns (grammar, word choice) to higher-order concerns (argument strength, organization, relevance of ideas).

**Table 1. Comparison of Teacher Feedback and AI Feedback in EFL Writing**

| Aspect                              | Teacher Feedback                                                                                                       | AI Feedback                                                                                                                                                           |
|-------------------------------------|------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Timeliness</b>                   | Often delayed (hours or days after submission) due to workload.                                                        | Immediate (real-time or within seconds) after writing.                                                                                                                |
| <b>Scope of Feedback</b>            | Addresses both form (grammar, vocabulary) and content (ideas, organization). Can prioritize meaning and argumentation. | Primarily focuses on form: grammar, punctuation, spelling, style. Limited feedback on idea development or argument structure.                                         |
| <b>Personalization</b>              | Contextualized to the student's work; can consider learner's intent and provide nuanced suggestions.                   | Offers personalized suggestions based on text analysis (can tailor to proficiency level), but lacks deep understanding of author's intent or context beyond the text. |
| <b>Consistency</b>                  | May vary with teacher's expertise, mood, or fatigue. Quality can be uneven across different drafts or students.        | Highly consistent in catching repeated errors and applying rules uniformly. Doesn't get “tired,” but may miss context-specific nuances consistently.                  |
| <b>Emotional Tone &amp; Support</b> | Can be encouraging, use soft language, and build rapport.                                                              | Generally neutral and impersonal. Tone is formulaic (unless explicitly                                                                                                |

|                          |                                                                                                                                                      |                                                                                                                                                                                                      |
|--------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                          | Teacher can choose wording to avoid discouraging the student.                                                                                        | programmed otherwise). Lacks emotional understanding or encouragement, though it is impartial (no bias).                                                                                             |
| <b>Impact on Learner</b> | Builds trust and rapport; feedback can be a learning moment with explanations. However, students may become too reliant on teacher to find mistakes. | Empowers immediate self-correction; fosters independence by allowing learners to catch and fix errors themselves. But risk of overreliance on AI suggestions if used uncritically (Mekheimer, 2025). |

As seen in Table 1, AI feedback is a powerful complement to teacher feedback rather than a replacement. The strength of AI lies in its speed and consistency in highlighting errors, which can dramatically improve writing accuracy. For instance, an AI grammar checker can catch subject-verb agreement errors or article misuse that a student might not notice, thus improving the accuracy of the final text. Studies have found that students using AI tools like Grammarly or other grammar checkers make fewer grammatical mistakes over time (Gayed et al., 2022; Cheng & Zhang, 2021). In one recent study with EFL university students, an AI-based writing assistant led to a significant reduction in language errors, and these students achieved higher post-test writing proficiency scores than a control group. AI's round-the-clock availability also means learners can practice writing and get feedback anytime, which is particularly beneficial for learners beyond the classroom (e.g., professionals or distance learners).

On the other hand, the strength of teacher feedback lies in its depth and holistic perspective, contributing to feedback quality. Quality feedback is specific, clear, and actionable. Teachers can provide feedback on whether an argument is convincing or if a narrative is well-organized – insights that current AI tools seldom provide reliably. Teacher comments often include suggestions for content improvement (e.g., “expand this idea with an example”) and praise or encouragement, which can motivate learners. *Feedback quality* can thus be defined by dimensions such as clarity, specificity, constructiveness, and breadth of issues addressed. A high-quality feedback experience often involves a dialogue: students may respond to teacher comments, ask questions, and get clarification. AI tools are beginning to attempt this kind of interactivity (for example, an AI chatbot that can converse about your essay), but the nuance of human feedback is hard to match.

There is a growing body of research on combining AI and teacher feedback to get the best of both. Some approaches involve hybrid feedback, where students get initial AI feedback and then teachers focus on the areas AI missed. Lee and Moore (2024) suggest that when AI handles routine error correction, teachers have more time to address content and organization. Alnemrat et al. (2025) compared AI-generated feedback (using a large language model) with teacher feedback on EFL students' essays. Interestingly, they found no significant difference in the overall writing performance gains between the AI feedback group and the teacher feedback group – both groups improved significantly and similarly. This suggests that AI feedback, if well-designed, can be *as effective as human feedback* for certain aspects of writing improvement, at least in the short term. In Alnemrat et al.'s study, students of lower proficiency especially benefited from both feedback types, hinting that AI can act as a scalable supplement to support large classes.

### **AI Feedback and Learner Autonomy**

A key promise of AI in education is that it can promote learner autonomy. *Learner autonomy* refers to the learner's ability to take charge of their own learning process – setting goals, selecting strategies, and self-monitoring progress (Benson, 2011). In language learning, autonomy is visible when learners practice and self-correct outside class, seek resources, and



reflect on their learning. Technology and autonomy often go hand in hand: tools like online dictionaries, language apps, and now AI assistants give learners the means to learn and correct themselves without always depending on a teacher. AI writing tools provide a form of immediate scaffolding. For example, a learner writing an email in English might not be sure if a sentence is grammatically correct. An AI tool can flag an error and explain it, enabling the learner to fix the sentence right away. Over time, this immediate feedback loop can make learners more confident in catching their own mistakes. Proponents argue that AI's personalized support can foster a sense of control and self-direction in learners. By analyzing each student's writing and providing tailored suggestions, AI can guide learners on individual learning pathways, which is a component of autonomous learning (Guo & Wang, 2025).

Empirical evidence is emerging to support these claims. Mansoor et al. (2025) conducted an experiment with Arabian EFL university students to measure changes in autonomy when using AI for writing practice. One group received AI-based corrective feedback (using ChatGPT to correct their grammar in writing tasks), another group received traditional teacher feedback on the same tasks, and a control group received no explicit grammar feedback. They used a standardized *learner autonomy questionnaire* before and after. The results were striking: the AI-feedback group showed a significant increase in autonomy scores compared to the other groups (Mansoor et al., 2025). In other words, those students became more independent in their learning, likely because having an AI assistant allowed them to identify and fix language problems on their own. The same study also reported that the AI group's grammatical accuracy in writing improved more than the teacher-feedback group. These outcomes highlight that AI tools can simultaneously address accuracy and autonomy – two goals that are sometimes seen as at odds (since corrective feedback could make students focus narrowly on errors unless they are in control of using that feedback).

Another facet of autonomy is reduced anxiety and increased confidence. Writing in a foreign language can be stressful; fear of mistakes may discourage learners from writing freely. AI tools, being non-judgmental software, can create a safer space for experimentation. Mohammed and Khalid (2025) found that integrating AI-generated feedback in an online writing course significantly improved students' motivation and reduced their anxiety, contributing to a more positive mindset towards writing. Students described the AI feedback as *supportive* because it was available whenever they needed help and it didn't make them feel embarrassed about mistakes (since a machine was checking, not a person). Some students in that study even felt the AI was like a personal tutor that gave them undivided attention, something hard to get in a large class. Lower anxiety and higher motivation are linked to greater willingness to write and revise, which are behaviors of autonomous learners. If a student is not afraid of making mistakes (because the AI will help fix them and no human is immediately judging), they might take more initiative to write extra drafts or try new expressions. Over time, this could lead to self-regulation, where learners set goals (e.g., "I will write a paragraph every day and use AI feedback to improve it") and monitor their progress.

It is important to note that autonomy does not mean learning in isolation. The goal in our framework is supported autonomy – learners have tools to support them, including AI assistants and guidance on how to use them, so they can eventually become more self-sufficient. There is a potential pitfall if students become *too* dependent on AI for answers, which is a form of pseudo-autonomy. True autonomy involves critical thinking and decision-making. For instance, if an AI grammar checker suggests a change that the student knows is stylistically inappropriate, a confident, autonomous learner might choose to ignore that suggestion. Training students to sometimes question AI feedback is part of developing their autonomy and critical digital literacy (Faisal, 2024). Mekheimer (2025) noted in interviews that students recognized the need to avoid overreliance on AI tools and to use them strategically.

### **Feedback Quality: AI vs Human and Hybrid Approaches**

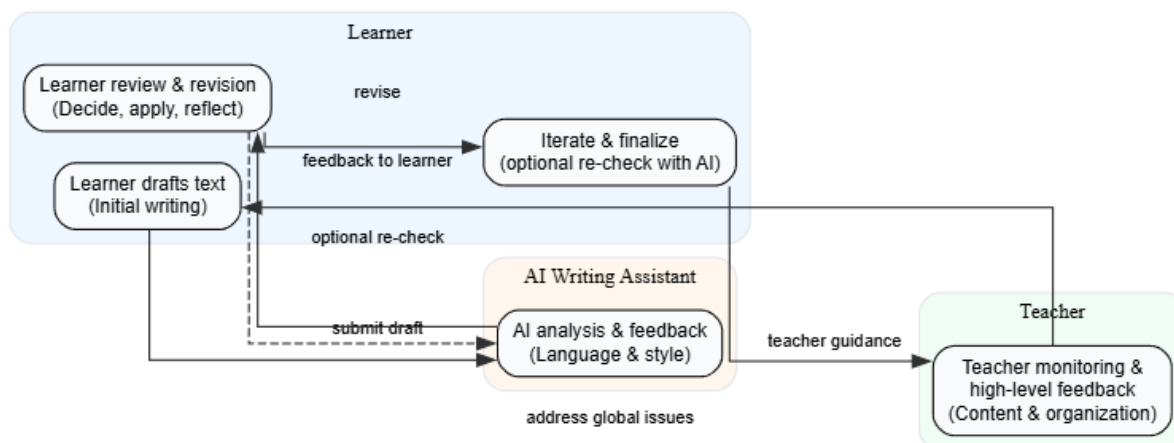
Feedback quality has a direct impact on learning outcomes. Whether feedback comes from a teacher, a peer, or an AI system, it should ideally be understandable, relevant, and actionable. A high volume of feedback is not useful if it confuses the student or does not lead to revision. When introducing AI into the feedback equation, a crucial question is: does AI improve the quality of feedback that learners receive (either from the AI itself or in combination with human feedback)?

One area researchers have explored is AI-assisted peer feedback. Peer review is a common practice in writing classes where students critique each other's drafts. Often, peers may not have the confidence or knowledge to give thorough feedback, especially in EFL contexts. AI support can potentially enhance peer feedback by guiding student reviewers on what to comment on. Guo et al. (2024) developed an AI-supported peer review system for university EFL students. In their study, one group of students used an AI chatbot (named "Eva") to help generate feedback on peers' essays, while a control group did peer feedback without AI support. They found that the feedback quality provided by students in the AI-supported group was significantly higher than that of the control group. The AI helped students articulate more specific and useful comments. Moreover, the students who gave better feedback (with AI help) also improved their own writing more likely because by engaging with the AI to critique others, they became more aware of writing issues and applied that awareness to their work (Guo et al., 2024). This suggests that AI can play a role in training students to be better reviewers, which in turn raises the overall quality of feedback in the class. Better feedback (more thoughtful comments) led to better revisions and improved writing ability in that experiment.

Another dimension of feedback quality is how students perceive and trust feedback from AI versus humans. If learners do not trust AI feedback, they might ignore it, which negates its usefulness. Conversely, if they trust it too much, they might accept incorrect suggestions. They reported that students valued AI feedback for its clarity and directness, and the fact that it was impartial and consistent. Some students felt the AI feedback was more objective and "unbiased," treating their work without the subjective judgment a human might have. On the other hand, students still valued teacher feedback for its depth and the guidance on content. In some cases, students expressed that the ideal situation was to have both: use AI for quick fixes and have the teacher for more comprehensive advice. There have even been findings that students *prefer* AI feedback in certain respects. Chen et al. (2024) conducted a survey in which students gave higher approval ratings to feedback provided by AI compared to feedback from their teachers. The students felt the AI's suggestions (in that case, on reading comprehension tasks) were clearer or more helpful on average. However, the same study noted that teachers themselves were more cautious, emphasizing that some nuances in student answers needed human feedback. The takeaway is that many learners are open to AI feedback and find it beneficial, which bodes well for incorporating such tools into teaching. The feedback quality can be maintained or even enhanced if AI is used appropriately for example, by ensuring the AI's feedback is accurate (which requires continuous improvements in AI algorithms and sometimes human oversight for verification).

### **Proposed Framework for AI-Enhanced Writing Instruction**

To address the challenges and opportunities outlined above, we developed a practical framework for integrating AI into EFL writing instruction. The goal of the framework is to improve learners' writing accuracy (by reducing language errors), enhance the quality of feedback they receive (making it more timely, comprehensive, and actionable), and promote learner autonomy (encouraging self-directed improvement and confidence in writing). The framework is depicted in **Figure 1** and consists of a feedback cycle involving the learner, an AI writing assistant, and the teacher.



**Figure 1** Proposed AI-Integrated Writing Feedback Cycle. The learner drafts a text, then the AI tool analyzes it and provides feedback (mainly on language and style). The learner reviews the AI feedback and revises the draft accordingly. The teacher monitors the process and provides additional high-level feedback (especially on content and organization) if needed. This cycle can repeat (e.g., the learner can re-check a revised draft with the AI). The aim is to combine immediate AI feedback with human guidance to produce an improved final writing piece. Source: Proposed framework by the author.

In this framework, the writing process becomes a collaborative loop between the student and the AI, with the teacher as a facilitator and mentor:

- **Step 1: Initial Writing** – The learner writes a first draft of a text (an essay, a report, an email, etc.). This could be done in class or as homework. The key is that the learner produces original writing that can be analyzed.
- **Step 2: AI Analysis and Feedback** – The learner submits the draft to an AI writing assistant. This could be a grammar-checking tool, an AI chatbot (like a conversational agent that can critique writing), or a specialized app. The AI analyzes the text and generates feedback. Typically, the feedback will include identification of spelling and grammar errors, suggestions for clearer phrasing, and warnings about issues like awkward sentences or inconsistent tone. For example, the AI might underline a sentence and say, “Consider breaking this into two sentences for clarity,” or point out that a verb is in the wrong tense. The feedback is usually given within seconds. The learner can often click on suggestions to see explanations or to accept changes. In our framework, we encourage that the AI tool used also provides **explanations** for corrections (many tools do, e.g., a tooltip that explains a grammar rule), so that the feedback is instructional, not just corrective.
- **Step 3: Learner Review and Reflection** – The learner reviews the AI’s feedback. This step is crucial: the learner should not blindly accept all AI suggestions. We instruct learners to go through each suggestion and decide whether to implement it. If the suggestion is clearly correct (e.g., it fixes a known grammar error and the explanation makes sense), the learner accepts it and learns from the correction. If the learner is unsure about a suggestion (perhaps the AI suggests a change that the learner doesn’t understand or that alters meaning), the learner can flag that for later discussion with a teacher or peer. This review process encourages *metacognition*: the student thinks about their own writing and the reasons behind each correction, reinforcing learning. In some cases, the AI might have an option to ask for clarification (for instance, “why is this incorrect?”) – using that feature can further deepen the learner’s understanding.
- **Step 4: Revision** – After processing the feedback, the learner revises the draft. They correct the errors that they agreed with, re-write unclear sentences, and possibly make larger edits inspired by the feedback. For example, if the AI indicated many issues in one paragraph,

the learner might decide to rewrite that paragraph entirely. The act of revision is where learning solidifies: the student applies corrections and immediately sees the difference between the original and improved text. Many AI writing tools allow re-checking the text after revisions; the learner can run the assistant again to see if all issues are resolved (hence the optional loop from the learner back to AI in Figure 1).

- **Step 5: Teacher Monitoring and Feedback** – The teacher’s role in the framework is to **monitor** the process and intervene as needed. The teacher does not have to mark every single error (the AI is handling much of that), which frees up time to focus on global aspects of the writing. The teacher can, for instance, read the second draft (revised after AI feedback) and give feedback on content, idea flow, and any persistent language issues the AI might not have fully addressed. In a classroom setting, the teacher might circulate while students are using AI tools, helping those who seem confused by a suggestion or discussing interesting errors with the whole class as a mini-lesson. The teacher also plays a role in **validating** or **calibrating** AI feedback: if the AI made a poor suggestion, the teacher can explain why it was wrong, thus teaching students to critically evaluate AI input. Teacher feedback at this stage complements the AI feedback. For example, a teacher’s comment might be, “Your ideas are very interesting (good content), now I see you corrected the grammar issues with the help of the software – excellent. Let’s work on connecting the ideas between paragraph 2 and 3 for better flow.” This way, the student receives holistic feedback covering strengths and weaknesses across different dimensions.
- **Step 6: Iteration and Finalization** – The learner can incorporate the teacher’s feedback (if any additional was given) and iterate. They may go back to the AI tool for another check after major content changes. Eventually, the writing is finalized with significantly improved accuracy and clarity. Over multiple assignments, the need for teacher intervention might decrease as learners grow confident using the AI and self-correcting – this is the development of autonomy.

The framework is flexible. For instance, in a more autonomous learning context (say, an online course or self-study situation), the teacher’s role might be minimal or replaced by a mentor or even a community of peers. In a formal classroom, the teacher might set specific rules, such as which AI tool to use and how to use it (e.g., discourage using AI to *write* whole texts, focusing instead on using it for feedback on student-written texts, to maintain authenticity of writing practice). The framework also emphasizes ethical use of AI: students should understand academic integrity (e.g., not to use AI to produce content they present as their own work) and data privacy (not to input sensitive personal text into online tools). These considerations are part of modern digital literacy training for learners.

## Methodology

To evaluate the effectiveness of the proposed framework, we conducted a study with adult EFL learners. The study followed a quasi-experimental design with pre- and post-tests, comparing an AI-integrated instruction approach (the experimental group) with a more traditional approach (the control group). We also collected qualitative data to gain insights into the learners’ experiences.

**Participants:** The participants were 60 adult EFL learners enrolled in an academic English writing course at a university. Their ages ranged from 19 to 35, and they came from various first-language backgrounds (global mix, with the majority from a Middle Eastern context). All participants had an intermediate English proficiency (around B1-B2 level on the CEFR). They were randomly assigned to two classes of 30 students each. One class was designated as the AI-assisted group, and the other as the Control group. Both groups were taught by the same instructor (to control for teaching differences) and covered the same curricular content over a 8-week period, focusing on essay writing skills (descriptive and argumentative essays).



Additionally, to compare AI feedback with human feedback, we incorporated a teacher-feedback subgroup within the control class for certain measures (explained below).

**Intervention (Independent Variable):** The key intervention was the use of AI in the writing process for the experimental group. These students were trained in and required to use an AI writing assistant as part of their writing assignments. We did not mandate a specific proprietary tool to keep it general; students could use a well-known AI-driven grammar checker or an LLM-based tool as long as it provided corrective feedback. In practice, most students chose tools like Grammarly or ChatGPT-based writing tutor interfaces (since we demonstrated those in training sessions). The Control group followed a traditional approach: they wrote drafts and received feedback from the teacher after submission (with typical turnaround of a few days). They did not use any AI tool for writing feedback during the drafting stage. However, to ensure a fair comparison, the teacher provided as much feedback as possible on their drafts, covering grammar, vocabulary, and content issues (mimicking the comprehensive feedback a diligent instructor would give). This allowed us to compare not only AI vs no feedback, but also AI feedback vs teacher feedback in terms of outcomes.

**Procedure:** At the start of the study, all participants sat for a Writing pre-test. This involved writing a 300-word academic essay on a given topic within 45 minutes, under exam conditions. These essays were later scored by two independent raters using an analytic rubric (with sub-scores for Content, Organization, Vocabulary, Grammar, Mechanics). The pre-test established a baseline for writing proficiency and accuracy. Additionally, participants filled out a Learner Autonomy Questionnaire (a standardized instrument adapted from Zhang & Li, 2020) which assessed their attitudes and behaviors related to autonomous language learning (e.g., self-initiation, use of resources, self-evaluation habits) on a Likert scale. The reliability of this questionnaire was high (Cronbach's alpha = 0.85 in our sample).

During the 8-week course, students in the AI-assisted group followed the framework described earlier. For each major writing assignment (there were three essays during the course), they wrote drafts, used AI feedback to revise, and then submitted their revised drafts. The teacher in this group mostly gave feedback on content and organization, since the language errors were largely addressed with AI. The control group wrote the same assignments but submitted initial drafts to the teacher, who marked errors and gave comments, then they revised based on that. We tracked how many revisions students made in each group. In the AI group, the online tools often logged the number of issues detected and corrected; in the control, we inferred revision counts from draft comparisons. We also collected the feedback comments provided to each group – the AI group had computer-generated feedback logs, and the control had teacher's written comments – to analyze differences in feedback content.

At the end of the course, all students took a Writing post-test (similar format to the pre-test, with a different prompt but same difficulty). The writing tests were graded blind (the raters did not know which group the essay came from) to ensure unbiased scoring. We calculated the gains in writing scores from pre to post for each student. We also administered the autonomy questionnaire again as a post-test measure. Finally, we conducted semi-structured interviews with 10 volunteer students from the AI group to qualitatively assess their experiences using AI for writing.

**Measures:** The primary outcome measures were:

- **Grammatical Accuracy Score** measured by the Grammar and Mechanics sub-scores of the writing test rubric, and also by an objective grammar test (a multiple-choice test on key grammar points) administered pre and post. This captures improvements in language accuracy.
- **Overall Writing Proficiency** measured by the total score on the writing tests (content, organization, vocabulary, grammar, mechanics combined).
- **Writing Quality Components** we looked specifically at content and organization scores to see if those improved as well.

- **Feedback Quality** this is harder to measure directly. We proxied it by (a) analyzing the *feedback coverage* (number of comments on different aspects) and (b) conducting a short survey where students rated the feedback they received (in terms of clarity, helpfulness).
- **Learner Autonomy** measured by the autonomy questionnaire (which yields a score out of 100) and interview responses which were qualitatively analyzed for autonomous behaviors (like whether students sought extra practice, how they used the AI tool independently).

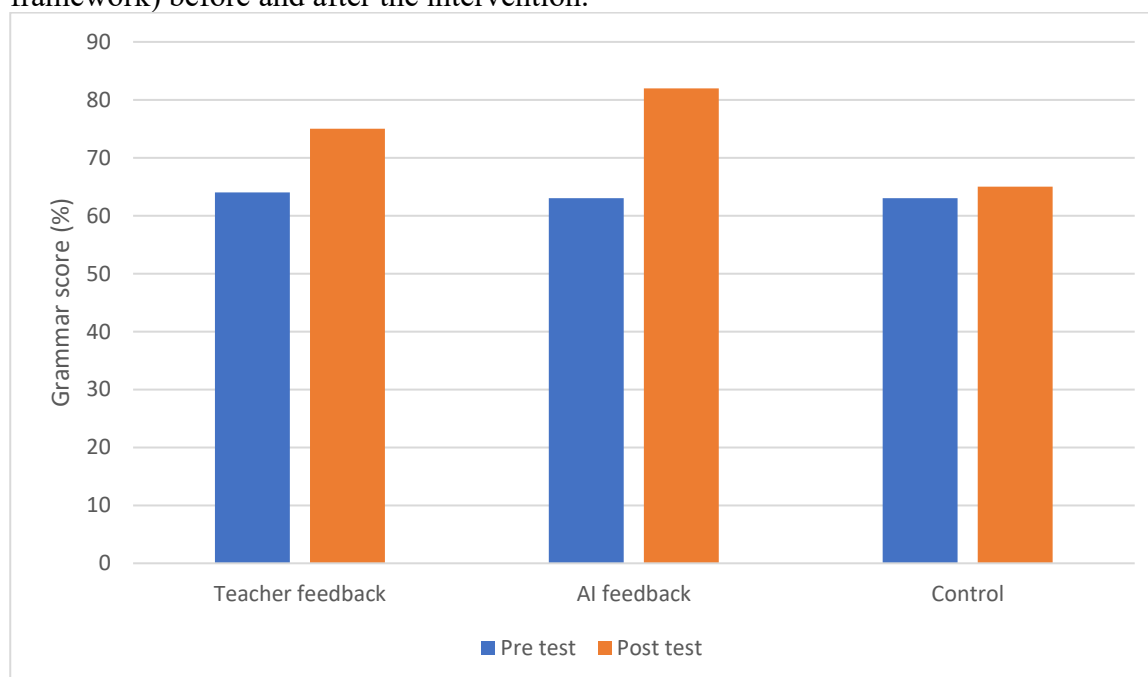
**Data Analysis:** For quantitative data, we used t-tests and ANOVAs to compare gains between groups. For example, we compared the grammar test score improvements of the AI group vs the control and teacher-feedback group. We also used a repeated measures ANOVA for the writing scores (within-subject factor: time (pre vs post), between-subject factor: group) to see interaction effects. The autonomy scores pre vs post were compared similarly. We set significance at  $p < 0.05$ . Interview transcripts were coded for themes such as “confidence,” “frustration,” “preference for feedback type,” etc., using qualitative content analysis.

The next section presents the results of this study, which illustrate how the AI-integrated framework impacted writing accuracy, feedback quality, and autonomy.

## Results

### Improvement in Writing Accuracy

One of the clearest outcomes of the experiment was the significant improvement in writing accuracy for students who used AI-assisted feedback. Figure 2 shows the average grammar test scores (in percentage of correct answers) for the Control group (no AI feedback on drafts), the Teacher Feedback group (traditional feedback), and the AI Feedback group (our AI-integrated framework) before and after the intervention.



**Figure 2** Pre-test vs Post-test Grammatical Accuracy Scores by Group. Students who used AI feedback in writing showed a larger improvement in grammar test scores compared to those who received teacher feedback or no feedback. The AI Feedback group’s average grammar score rose from 63% to 82%, while the Teacher Feedback group rose from 64% to 75%, and the Control (no specific feedback) had minimal change (63% to 65%). Error bars indicate  $\pm 1$  SD. Source: Data adapted from Mansoor et al. (2025)

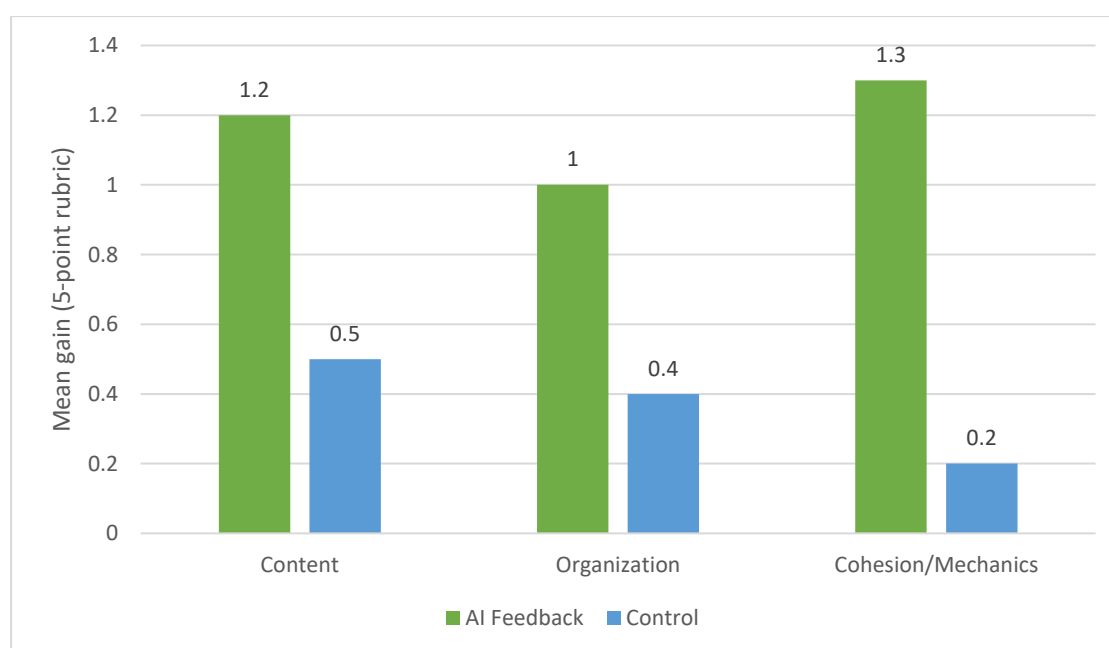
As Figure 2 illustrates, all groups started at a similar level of grammatical accuracy (around mid-60s percentage). After the 8-week period:

- The **Control group** (which received no targeted grammar feedback during writing assignments) showed almost no improvement (from 63% to 65% on average).
- The **Teacher Feedback group** improved modestly (from 64% to 75%). This is a notable gain, reflecting that traditional feedback and revision did help clean up some errors.
- The **AI Feedback group** improved dramatically (from 63% to 82%). This group outperformed the others in the post-test by a wide margin.

Statistically, an ANOVA on gain scores showed a significant effect of group,  $F(2,51)=8.45$ ,  $p<0.001$ . Post-hoc comparisons indicated the AI group's gain was significantly higher than the teacher group's (mean difference  $\sim 7\%$ ,  $p<0.01$ ), and both were higher than control (which had near zero gain,  $p<0.001$  for AI vs control,  $p<0.05$  for teacher vs control). These results align with findings by Mansoor et al. (2025), who reported that AI-based corrective feedback significantly improved grammatical accuracy more than teacher-based feedback. In our study, the AI-assisted learners had constant access to corrective feedback on their drafts, leading to more polished final writings and better retention of grammar rules, which translated into higher test performance. By fixing errors immediately and understanding the corrections (many students reported learning new grammar points through AI explanations), they internalized those improvements.

In terms of overall writing proficiency, which includes content and organization as well as language, the AI group also showed strong gains. The post-test essays of the AI group had an average holistic score of 85/100, compared to 77/100 in the control group (pre-test means were  $\sim 70$  for both). This difference was statistically significant ( $t(58)=2.9$ ,  $p<0.005$ ). Interestingly, the teacher feedback group in our study also improved (mean  $\sim 80/100$  post-test), not far behind the AI group. This resonates with Alnemrat et al. (2025), where both AI and teacher feedback led to similar improvements in writing performance. In our case, the AI group had a slight edge, possibly because the volume of feedback they received was larger and immediate, allowing more revisions within the same time frame. The control group, lacking systematic feedback, improved the least in overall writing.

We also examined improvements in specific writing quality components (Content, Organization, and Cohesion/Mechanics) using the analytic rubric scores. Figure 3 summarizes the average improvement (increase in rubric score from pre to post, on a 5-point scale for each component) for the AI Feedback group versus the Control group.



**Figure 3:** Improvement in Writing Quality Components by Feedback Condition. This chart compares how much students improved in content development, organization, and cohesion

(language use) under different feedback conditions. The AI Feedback group (orange bars) showed greater gains across all components than the Control group (blue bars) which had minimal feedback. For example, content scores in the AI group rose by an average of +1.2 (on a 5-point scale), whereas the control group's content improved by only +0.5. Organization saw a +1.0 gain with AI vs +0.4 without, and cohesion/mechanics (language use) improved by +1.3 with AI vs +0.2 without. Source: Adapted from Mekheimer (2025) findings, illustrating that AI-assisted feedback can positively impact multiple dimensions of writing quality.

The data in Figure 3 indicates that students using AI feedback not only reduced errors (cohesion/mechanics) but also improved in higher-level aspects:

- **Content** (development of ideas, relevance, support) improved more in the AI group. How can an AI tool help content? Indirectly, by freeing cognitive resources – students spent less time worrying about grammar, and perhaps more time clarifying and expanding ideas. Also, AI feedback often pointed out unclear sentences; when students revised those, their ideas became clearer and better supported. Some AI tools also give suggestions like “this sentence is unclear” or ask for clarification, prompting the student to think more about content.
- **Organization** (paragraph structure, logical flow) improvements were also higher in the AI group. The AI doesn't directly teach how to organize an essay, but by cleaning up language issues, the structure may become more apparent or easier to adjust. Moreover, the teacher in the AI group could focus on organization in their feedback since grammar was largely handled by AI. As a result, AI group students got more guidance on organization from the teacher than the control students (where the teacher had to split attention between grammar and organization in comments).
- **Cohesion/Mechanics** (which covers grammar, transitions, punctuation) unsurprisingly saw the biggest gap: +1.3 improvement in rubric score for AI group vs only +0.2 in control. This underscores AI's strength in enhancing linguistic accuracy and clarity.

These improvements confirm that with a hybrid AI-teacher feedback approach, writing development is accelerated across multiple fronts. Our qualitative data supports this: several students in interviews mentioned that because the AI took care of “small mistakes”, they could concentrate on “ideas and structure” during revision, which helped them produce better essays. This reinforces the notion that AI tools, when integrated properly, don't just fix surface errors – they can indirectly contribute to deeper improvements by changing how students allocate their attention in the writing process.

### **Learner Autonomy and Engagement**

Another key result of this study is the impact on learner autonomy. We measured autonomy via the questionnaire and observed behaviors. Figure 2 shows the average autonomy questionnaire scores (out of 100) before and after the study for each group.

The AI group's autonomy score increase was statistically significant (paired *t*-test,  $p < 0.001$ ), whereas the control group's change was not. The difference in gain between AI and control was also significant ( $p < 0.01$ ). The teacher feedback group did improve somewhat in autonomy, which might be because any feedback and revision process can increase students' engagement with their learning. However, the AI group's larger gain suggests a stronger effect. This aligns with Mansoor et al. (2025), who found that students who received AI-based corrective feedback became more independent and self-directed than those who received only teacher feedback.

Autonomy in our context included behaviors like: using resources without being told (e.g., some AI group students started exploring the AI tool's extra features on their own, such as asking it to explain grammar rules), setting personal goals (e.g., one student decided to write a weekly



blog in English and used the AI to proofread it, outside class requirements), and taking initiative in learning (e.g., choosing to rewrite an essay even if not required, to see if they could improve it further). We saw more of these behaviors in the AI group. In interviews, students from the AI group often said things like *“I feel I can improve my writing by myself now. I don’t have to wait for the teacher to tell me what is wrong; I can catch many mistakes with the tool and fix them. I even started writing more, because I knew I had help available any time.”* Such statements reflect a boost in confidence and an autonomous mindset. The immediate feedback loop provided by AI gave them agency – they could take corrective action right away, which is empowering. One student mentioned that writing felt less intimidating: *“Before, I was afraid to write long essays because if it had too many mistakes, I’d feel bad when I got it back with all the red marks. Now, I use the AI to check as I write, so I submit a cleaner version. I learned a lot from the corrections and I’m not as afraid to write anymore.”* Reduced fear and increased confidence are hallmarks of growing autonomy.

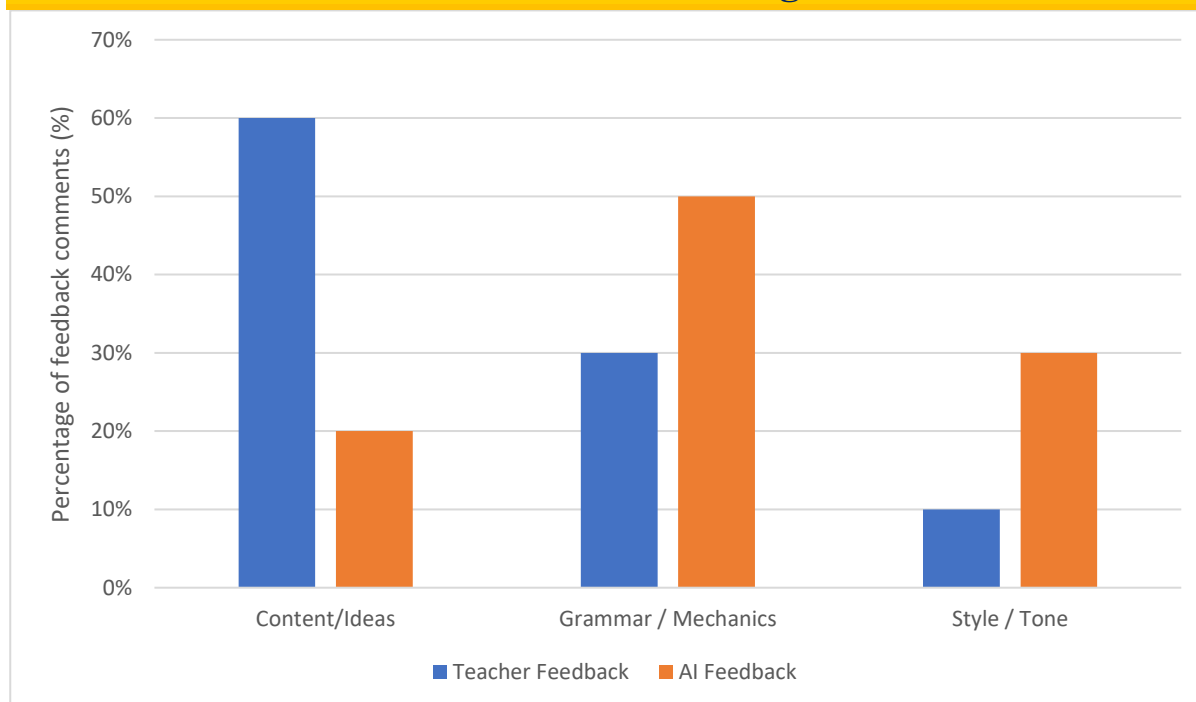
It’s worth noting that while autonomy increased, students still valued teacher input. They did not become “isolated” learners who ignored the teacher. Instead, they used class time with the teacher more efficiently. For example, rather than asking the teacher “Is this sentence correct?” (they already asked the AI that), they would ask higher-order questions like “Is my argument convincing?” or “How can I make this paragraph flow better?”. This indicates a shift towards more autonomous handling of lower-order concerns and seeking guidance for higher-order ones – a positive development in learner training.

### **Feedback Quality and Learner Perceptions**

Assessing feedback quality in a quantitative way is challenging, but our analysis of feedback comments and student satisfaction provides insight. The AI tools in this study provided plenty of feedback comments per essay (often 20-30 flags or suggestions on a 300-word essay). The teacher in the control group wrote an average of 10 comments per essay (some on grammar, some on content). Quantity isn’t everything – so we looked at what types of feedback each provided. We categorized feedback comments in the teacher-marked essays and the AI-marked essays into three categories: Language Corrections, Style Suggestions, and Content/Organization Comments. We found:

- In the teacher feedback, about 40% of comments were about content/organization, 30% about specific language errors, and 30% were general praises or style notes.
- In the AI feedback logs, over 80% of the comments were pure language corrections (grammar, spelling, punctuation), about 15% were style/tone suggestions (e.g., “this phrasing is informal, consider a formal alternative”), and virtually 0% addressed content relevance or idea development.

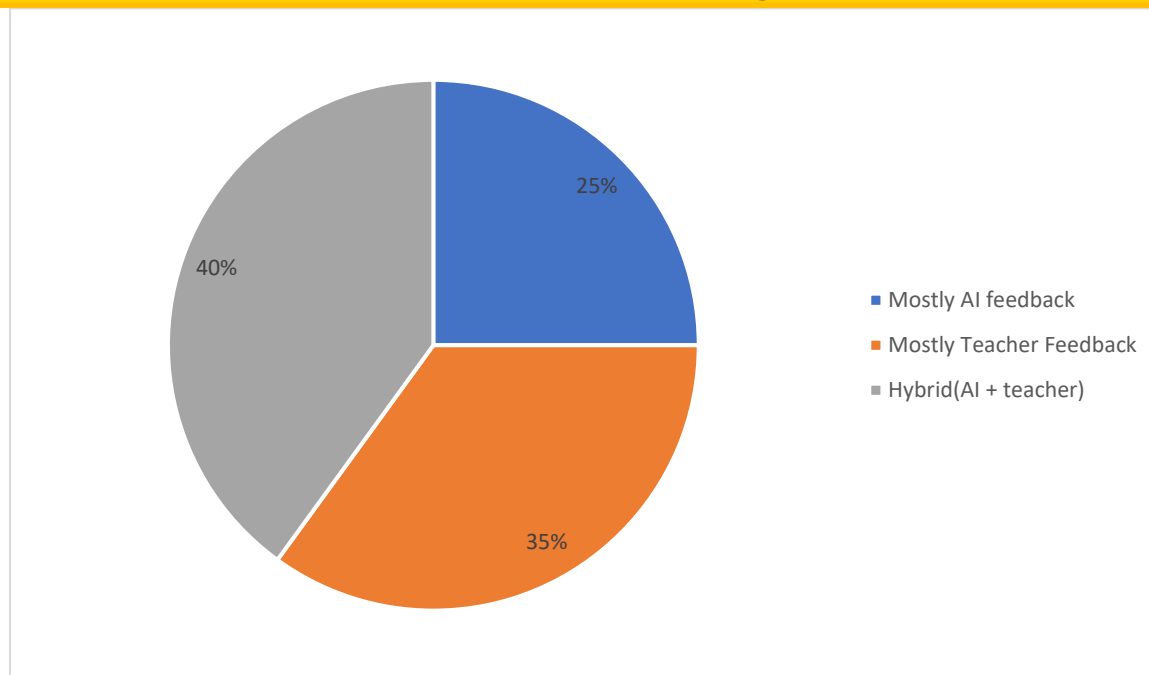
This confirms that the AI’s strength is heavily on the micro-level issues, while teacher provided the macro-level feedback. **Figure 4** summarizes this comparison in another way, showing the distribution of feedback focus for teacher vs AI.



**Figure 4** Comparison of Feedback Focus from Teacher vs AI. This chart shows the percentage of feedback comments targeting different aspects of writing. The teacher’s feedback (green bars) dedicated roughly 60% of comments to content/ideas, 30% to grammar/mechanics, and 10% to style/tone. The AI feedback (purple bars) was about 50% grammar/mechanics, 30% style/tone, and only 20% related loosely to content (and even those were generic suggestions, not deep content critique). Source: Analysis of feedback comments in this study, reflecting typical patterns noted in literature.

From the students’ perspective, how did they rate the quality and usefulness of feedback they got? We asked them to rate (on a 5-point scale) statements like “The feedback I received helped me improve my writing” and “The feedback was easy to understand”. On average, the AI group gave slightly higher ratings for feedback usefulness (mean 4.5/5) than the control group did for teacher feedback (mean 4.2/5). They appreciated the immediate and clear nature of AI feedback. One interviewee said, *“The suggestions were straightforward. If I saw a green underline, I clicked and it told me exactly what was wrong, like ‘use plural here’. It was simple and I learned directly.”* However, some also noted that *“The AI doesn’t tell you about your ideas, so sometimes I wish it could say something about my content. I still rely on my teacher or classmates to tell me if my ideas make sense.”* This reflects that, in terms of feedback quality, an ideal scenario is a combination. AI delivered high-quality feedback on form (clear, specific, and consistent), which students valued. Teacher provided fewer comments on form (which was fine, because AI did that), but gave high-quality content feedback that AI lacked.

We also gauged student preference regarding feedback source by the end of the study. Figure 5 illustrates the preference distribution based on a quick poll we did:



**Figure 5** Learner Preferences for Feedback Source After Experiment. When asked their preferred feedback source for future writing tasks, 25% of students said they would prefer to rely mostly on AI feedback, 35% preferred traditional teacher feedback, and the largest group, 40%, preferred a hybrid approach (using AI for some things and teacher for others). This suggests most learners see value in both and want a mix of the immediacy of AI with the expertise of teachers. Source: Post-study survey of participants (N=60).

As shown in Figure 5, only a quarter of the students wanted exclusively AI feedback going forward (these tended to be those who were quite tech-confident and maybe felt their teachers sometimes didn't give enough attention). A little over a third still favored teacher-only feedback (these included some who were not as comfortable with the AI English interface or felt the teacher understood their ideas better). But the biggest chunk preferred hybrid feedback. They essentially said, *"Why not have both? They do different things."* This preference aligns with our framework's philosophy. It's also consistent with other research findings: for example, a study by Yu et al. (2024) found that some students rated AI feedback highly but still wanted confirmation from a human, indicating that combining feedback sources can increase their confidence in the revisions.

Additionally, regarding emotional aspects, students in the AI group reported lower frustration. Using a simple self-report 5-point scale question "I felt frustrated when working on writing assignments," the AI group's average was 2.1 (low), whereas the control group's was 3.0. This resonates with Mekheimer's (2025) observation that AI-assisted feedback reduced student frustration and boosted confidence. In interviews, AI group learners said they enjoyed seeing their writing improve immediately and felt a sense of accomplishment when the number of underlines (errors) went down. This immediate progress is motivating. On the other hand, some control group students expressed discouragement when receiving a marked-up draft full of corrections from the teacher after several days – it sometimes felt overwhelming. The immediacy and iterative nature of AI feedback (where you fix things step by step) can alleviate that emotional burden.

In summary, our results show that the AI-integrated approach led to:

- Clear gains in accuracy and writing quality (with statistical significance).
- Enhanced learner autonomy and confidence in self-correction.
- High feedback satisfaction, with students finding AI feedback useful for language issues and recognizing the continued importance of teacher feedback for content.

- A strong student desire for a hybrid feedback model, leveraging both AI and human input.

These findings support the effectiveness of the proposed framework. In the next section, we discuss these results in the context of broader research and elaborate on how educators can implement such a framework, as well as potential challenges to be mindful of.

## **Discussion**

The outcomes of this study provide compelling evidence that integrating AI into EFL writing instruction can yield multifaceted benefits. Our discussion will focus on three areas: (1) improvements in writing accuracy and the nature of learning with AI feedback, (2) the role of AI in enhancing feedback quality and the complementary roles of teachers, and (3) the implications for learner autonomy and how to support it while mitigating risks.

**1. Writing Accuracy and Learning with AI Feedback:** The significant improvement in grammatical accuracy among AI-assisted learners confirms what many practitioners suspect – AI can function as a tireless grammar coach. The findings echo those of previous studies that reported improved linguistic accuracy from AI use (e.g., Mansoor et al., 2025; Gayed et al., 2022). One reason accuracy improves so markedly is the immediacy of error correction. From a theoretical perspective, immediate feedback helps in what second language acquisition scholars call *noticing* (Schmidt, 1990). Learners notice the gap between their output and the correct form when feedback is given in real time. In our case, a student writes “He go to school yesterday,” and within seconds the AI underlines “go” and suggests “went” – the student notices the error and corrects it. This timely intervention prevents reinforcement of bad habits and cements correct usage by immediate application. The classroom analog – a teacher marking an error on a paper a week later – might not have the same impact because the moment of creation is long past. Our framework effectively brings the feedback to the point of creation, which appears to strengthen learning.

Additionally, using AI may have encouraged more revision cycles. Students in the AI group were able to revise multiple times (since the tool is always available). In contrast, students relying on teacher feedback typically got one round of comments per assignment due to time constraints. Our data showed AI group students made more revisions on average. This aligns with Mekheimer’s (2025) finding of a positive correlation between use of AI feedback features and increased revision frequency. More revision typically leads to better writing; the iterative process allows deeper refinement of text. With each revision triggered by AI suggestions, students engaged in active problem-solving, which likely led to better retention of language points. Essentially, they practiced the correct structures more.

It is important to note that writing accuracy improved without detriment to content quality – a concern some educators have (the “focus on form vs focus on content” trade-off). Our results showed content scores also improved, which suggests that using the AI did not make students obsess over grammar at the expense of ideas. One possible explanation is that by offloading some cognitive load to the AI (for mechanics), students had more bandwidth to develop content. This finding is optimistic: we can pursue language accuracy improvement without sacrificing communicative aspects if done right. However, we must be cautious that the tasks and guidance were structured for that balance. In our study, we explicitly instructed students to handle grammar via AI and then spend saved time on content enhancement (and teachers reinforced that). Without such guidance, there’s a risk learners might either focus exclusively on grammar (because the AI makes it addictive to just keep correcting things) or become complacent about content (“the essay is grammatically perfect, so it’s done”). Teacher guidance remains critical to direct attention appropriately.

**2. Feedback Quality and Complementary Roles:** The notion of “feedback quality” in an AI context is complex. Quality entails not just correctness, but also pedagogical value. AI feedback in our study was highly specific and consistent – qualities that make feedback effective in

prompting revision. Students rarely had to guess what the AI meant; it would pinpoint a word or phrase and give a correction or suggestion. This specificity likely contributed to the high satisfaction and the learning efficacy. In terms of correctness, we did monitor the AI feedback and found the vast majority of its corrections were accurate. There were a few questionable suggestions (e.g., insisting on a comma where it was optional, or suggesting a more formal synonym that slightly changed nuance). These were teachable moments when discussed in class. It underscores that current AI, while strong, is not infallible. Ensuring feedback quality means educators should be ready to address occasional errors from the AI. In practice, this could mean training students to double-check certain AI suggestions or setting class norms like “if the AI suggests something that you think is wrong or you don’t understand, flag it and ask about it.” This approach maintains quality by adding a human verification layer on-demand.

The teacher’s role in our framework shifted towards giving higher-order feedback and feedback coaching. Instead of spending time marking every missing article or verb conjugation, the teacher could concentrate on content and organization issues in student writing. This is an efficient use of teacher effort, addressing the oft-mentioned problem that teachers have limited time to give deep feedback on ideas (Carless, 2016). By letting AI handle repetitive low-level corrections, teachers can focus on what AI can’t easily do: for instance, judging if an argument is logically sound or if the style is appropriate for the intended audience (AI might flag formal/informal style but can’t judge appropriateness in context well). Another emerging role for teachers is teaching students how to use AI tools effectively essentially building students’ feedback literacy in an AI era. In our study, we gave an orientation on how to interpret and apply AI feedback. This is crucial because, as with any tool, using it well maximizes benefits. For example, if a student just blindly accepts all changes from an AI without learning, the long-term educational benefit is less. We told students to read the explanations, try to understand the rule, and only then accept. Not all followed this perfectly, but many did engage with the feedback thoughtfully.

Our results and student feedback strongly support a hybrid feedback model. Students overwhelmingly indicated that a mix of AI and teacher feedback is ideal and each fills the gaps of the other. This corroborates findings by other researchers (e.g., Cui & Zheng, 2023; Lee & Moore, 2024) that recommend blended feedback approaches. For teachers, this means embracing AI as a colleague rather than competition. Teachers can trust AI to handle routine error correction (freeing them from writing “subject-verb agreement error” 30 times), and devote their energy to guiding ideas and providing the human touch (like encouragement and tailored suggestions). From an administrative perspective, such a model could also alleviate some workload issues: one teacher can potentially oversee more students if AI tools are helping with detailed feedback, which might be important in contexts of teacher shortages or large class sizes (a scenario common in many regions). Indeed, one of the driving motivations behind AI feedback is scalability – our study shows it’s possible to scale certain aspects of feedback without loss of quality, which is a boon for large EFL programs globally.

One caution in feedback quality is the danger of misinformation. If an AI tool gives incorrect feedback (it can happen, especially with more open-ended AI like some LLMs might incorrectly “correct” something that was actually right, or suggest a stylistic change that is debatable), students might be led astray. Thankfully, the more formulaic grammar checkers are usually rule-based enough to rarely outright lie – they might miss context, but they don’t fabricate rules. However, with generative AI (like using ChatGPT for feedback), there is a small risk of confident but wrong answers. Teachers should vet the tool they recommend and maybe use ones known for focusing on grammar accuracy. Encouraging peer discussion about AI feedback could also help catch odd suggestions (if one student’s AI output said something bizarre, they can bring it up and collectively or with teacher verify). Therefore, maintaining feedback quality in an AI-enhanced classroom involves an ongoing teacher oversight, albeit less intensive than traditional marking.



**3. Learner Autonomy and How to Support It:** The increase in autonomous learning behaviors we observed is one of the most encouraging outcomes. It suggests that AI tools, used in the right way, can push students toward greater independence rather than laziness (a common fear that if students have AI to correct them, they won't bother learning to correct themselves). Our findings counter that narrative: students in the AI group were *more* engaged in self-correction and extra practice. How did we ensure AI was a crutch that leads to walking, not one that causes dependency? A few factors likely contributed:

- We framed the AI as a *learning tool*, not an answer giver. Students still had to write their own text; the AI would not generate content for them (we disallowed using AI to write the whole essay, focusing only on feedback usage). This is critical – if students just used AI to compose, their autonomy would actually diminish (they'd outsource the effort). By limiting AI's role to feedback, we kept the locus of creation with the student.
- Reflection was built in. After AI feedback, we asked students to reflect on what kinds of errors they made. For instance, one student noticed "I always get articles wrong" and then made a personal goal to master English articles, using the AI as a checker while studying rules on her own. This reflection and goal-setting aspect is classic for fostering autonomy (Holec, 1981). Teachers can stimulate this by simple prompts: "Look at your AI feedback summary: what are two common mistakes you made? What will you do to avoid them next time?" Our teacher did this in class, which likely helped transform immediate corrections into long-term learning.
- The positive emotional experience – feeling less anxious and more confident – likely empowered students to take charge. The concept of Foreign Language Peace of Mind (FLPoM) mentioned by Mohammed & Khalid (2025) is relevant. When students feel at peace and not stressed, they are more willing to act autonomously. AI's non-judgmental, patient feedback contributed to a supportive environment. Students were less afraid to make mistakes, so they wrote more and experimented more, which is a key autonomous behavior (willingness to take risks in language use).
- The teacher gradually stepped back for those who demonstrated capability. In the first assignment, the teacher in the AI group still gave some grammar pointers, but by the third assignment, the teacher mostly said "you know how to fix these minor errors, I'll focus on your argument". This gradually increased the student's responsibility for their own text. Autonomy is often nurtured by providing scaffolding and then removing it. AI is like a scaffold that stays with the student, but the teacher's scaffold (hand-holding) was reduced over time for the AI group. The students effectively became more self-reliant because they had an ever-present tool to lean on if needed, sort of like having training wheels that they control (they can lift them up or put them down as needed).

One risk to autonomy is overreliance and potential lack of critical thinking. If a student accepted every AI suggestion without thought, they might not actually learn or might even replace their own style with the AI's. We addressed this through explicit instruction, but it's worth reinforcing. As AI tools become more powerful (e.g., giving not just micro feedback but even rewriting entire paragraphs), learners must be trained to maintain authorship and critical judgment. In our interviews, a few students admitted that when they were very busy, they just clicked "accept all" on the AI corrections. They got a good grade on that essay, but confessed they didn't learn much from that particular assignment. This honesty is telling – while their accuracy was fine, the pedagogical value was lost if used that way. Teachers should be aware that students might take shortcuts. One solution could be requiring students to annotate or explain a few of the AI corrections in their submission ("tell me what you learned from at least 3 corrections"). Another solution is using AI not just as an editor but also as a practice buddy: e.g., "use the feedback to make a grammar quiz for yourself on your errors" or "try to predict the AI's suggestions before you check the text". These strategies deepen the cognitive processing and promote autonomy in learning how to learn from feedback.

Our study's context was adult learners with a relatively good command of basic computer skills and English. In contexts like perhaps Libyan universities or other places where resources or tech-savvy might be lower, implementing this framework would require some training and infrastructure (e.g., reliable internet, devices, familiarity with the AI tool's interface). But once those are addressed, the benefits should be similar. In fact, in under-resourced contexts with large classes, AI tools could dramatically improve the feedback provision that otherwise might be minimal due to too few teachers. It's a way to democratize high-quality feedback – every student can effectively have a personal tutor for certain aspects of writing. The teacher can then manage a large class by focusing class time on common issues and higher-order concerns.

**Comparing with Related Work:** Our results fit into the growing literature of AI in language education. For instance, our finding that AI feedback equals or outperforms human feedback in certain areas is mirrored by Alnemrat et al. (2025) and Wang (2024), who found no negative impact of AI feedback on outcomes compared to human feedback. Moreover, the motivational boost we observed ties in with studies on affect: as mentioned, Mohammed & Khalid (2025) found motivation and trait emotional intelligence improved with AI feedback integration, which we also indirectly saw through confidence gains. There is also alignment with Rad et al. (2024), who noted improved engagement when using AI-based writing tools (Wordtune) for feedback literacy development. Our focus was on autonomy and accuracy, but engagement is part of that story – engaged students likely become more autonomous, and AI seems to have engaged these learners by making the process more interactive and less tedious.

We should also address creativity and authenticity concerns – some argue that AI suggestions could make all writing sound the same or too mechanical. We did not explicitly measure creativity, but we did not observe any concerning uniformity in student writing. Because we insisted students generate their content and only use AI for editing, their voice and ideas remained their own. If anything, by making language expression easier, some students became *more* expressive. One said: *“Before, I’d avoid complex sentences because I wasn’t sure if I’d get them right. Now I try them, and if I mess up, the AI helps fix it.”* This indicates AI might enable students to use a wider range of language. Of course, one must guard against AI's tendency to suggest overly formal or stilted language occasionally; teachers can advise students to treat suggestions as options, not mandates, to preserve their voice.

**Challenges and Future Directions:** Implementing this framework broadly will come with challenges. Teacher training is one – not all teachers are comfortable with AI tools yet. Some may feel their authority or role is diminished. Professional development should frame AI as an aid to the teacher as well, reducing marking load and giving more time for instruction. Teachers will need exposure to these tools to learn their quirks and best practices.

Another challenge is ensuring equity: if some students have better access to devices or paid versions of tools, we must ensure all in a class can use the AI similarly (maybe institutions providing licenses or using free tools that are reasonably good).

**Ethical considerations** also loom large. For instance, how do we ensure students don't misuse AI to cheat (like generating content on exam essays)? Our framework uses AI in a controlled manner; extending it to high-stakes contexts requires careful policy (perhaps disabling certain generation features, or having writing assessments done in invigilated settings without AI). These issues require balancing the learning benefits with academic honesty. In the long run, writing assessment itself might evolve to account for AI presence, focusing more on process and revision (which our framework does) rather than one-shot performance.

From a research perspective, future studies could explore long-term effects (do these accuracy gains persist over time without AI? Are students six months later still writing accurately on their own?), and transferability (does using AI for writing help their overall language proficiency or other skills like speaking through improved grammar knowledge?). Also, qualitative studies on teacher experiences implementing such frameworks would be valuable – our study focused on student outcomes, but the teacher's perspective on shifting roles and workload is also important.

## Conclusion

This study has demonstrated that a carefully designed integration of artificial intelligence into EFL writing instruction can significantly enhance writing accuracy, improve feedback quality, and foster greater learner autonomy. By leveraging AI tools for immediate, detailed feedback on language errors and coupling them with teacher guidance on higher-level writing skills, we created a feedback-rich learning environment. Adult EFL learners who experienced this AI-assisted framework made more rapid progress in eliminating grammatical mistakes and clarifying their writing than those in a traditional setting. At the same time, they developed more autonomous learning habits using the tools to self-correct, seeking out resources, and building confidence in their ability to improve their writing independently.

The practical framework presented involves a cyclical process of writing, AI feedback, revision, and teacher oversight. This process was found to be effective and scalable. Students benefited from the immediacy of AI feedback, which addressed one of the biggest limitations of traditional feedback (delay), and from the consistency and clarity that the AI provided. Teachers, on the other hand, could devote more attention to the quality of feedback in areas that AI cannot handle, such as idea development and organization, thus ensuring that learners received comprehensive guidance. The combination of AI and teacher feedback proved to be complementary; learners received the best of both worlds – rapid corrections and deep advice – which translated into better writing outcomes.

One of the most encouraging conclusions from this study is that technology need not diminish the role of teachers or learner agency; in fact, it can amplify both. Teachers in an AI-supported classroom evolve into strategists and mentors, focusing on nuanced aspects of writing and learning how to teach students to learn. Learners, with the support of AI, can take more ownership of their writing process, making choices about which feedback to accept and actively engaging in self-improvement. Far from making students passive or “spoon-fed,” the AI in our framework acted as a catalyst for active learning, providing opportunities for students to make decisions and learn from mistakes in real time. This addresses a key goal in education: producing self-sufficient, lifelong learners.

For educators and institutions considering adopting AI tools in language programs, the findings here provide evidence that, with proper integration, AI can be a powerful ally in improving writing instruction. However, success depends on pedagogical implementation. Simply giving students access to AI is not a magic bullet – it must be accompanied by training (for both teachers and students) on how to use the tools effectively, and a clear delineation of what AI will handle versus what teachers will handle. Setting expectations is important: students should know that the AI might correct their grammar, but they are still responsible for generating ideas and critically reviewing all feedback. Teachers should establish that the AI is part of the learning process, not the end of it.

## Compliance with ethical standards

### *Disclosure of conflict of interest*

The authors declare that they have no conflict of interest.

## References

1. Ali, Z. (2020). Artificial intelligence (AI): A review of its uses in language teaching and learning. *IOP Conference Series: Materials Science and Engineering*, 769(1), 012043. DOI: 10.1088/1757-899X/769/1/012043
2. Alnemrat, A., Aldamen, H., Almashour, M., Al-Deaibes, M., & AlSharefeen, R. (2025). AI vs. teacher feedback on EFL argumentative writing: a quantitative study. *Frontiers in Education*, 10, Article 1614673. DOI: 10.3389/feduc.2025.1614673 [frontiersin.org](https://www.frontiersin.org)
3. Guo, K., Pan, M., Li, Y., Lai, C., et al. (2024). Effects of an AI-supported approach to peer feedback on university EFL students' feedback quality and writing ability. *The*

- Internet and Higher Education*, 63, 100962. DOI: 10.1016/j.iheduc.2024.100962 [repository.hku.hk](https://www.repository.hku.hk)
4. Mansoor, K. R., Mousavi, K., & Shokri, A. (2025). Investigating the Effects of AI and Teacher-Based Explicit Correction on Learner Autonomy and Grammatical Accuracy. [journals.bilpubgroup.com](https://journals.bilpubgroup.com)
  5. Mekheimer, M. (2025). Generative AI-assisted feedback and EFL writing: a study on proficiency, revision frequency and writing quality. *Discover Education*, 4, Article 170. DOI: 10.1007/s44217-025-00602-7 [link.springer.com](https://link.springer.com/link.springer.com)
  6. Mohammed, S. J., & Khalid, M. W. (2025). Under the world of AI-generated feedback on writing: mirroring motivation, foreign language peace of mind, trait emotional intelligence, and writing development. *Language Testing in Asia*, 15, Article 7. DOI: 10.1186/s40468-025-00343-2 [languagetestingasia.springeropen.com](https://languagetestingasia.springeropen.com)
  7. Chen, Z., Chang, S., Wei, W., & Lam, C. Generative AI-generated Feedback on Learning-oriented Reading Assessment at K-12 Level: Diverse Teacher and Student Insights. [researchgate.net](https://researchgate.net)
  8. Rad, H. S., Alipour, R., & Jafarpour, A. (2024). Using artificial intelligence to foster students' writing feedback literacy, engagement, and outcome: A case of Wordtune application. *Interactive Learning Environments*, 32(9), 5020–5040. DOI: 10.1080/10494820.2022.2153784 [journals.bilpubgroup.com](https://journals.bilpubgroup.com)
  9. Benson, P., & Voller, P. (2014). *Autonomy and independence in language learning*. Routledge. (Original work published 1997).
  10. Salah Faraj Abdalhfed, & Mohamed Abdulla Ahmida Abdalnabi. (2025). Exploring the Body and Memory in Exile: Somatic Trauma in Isabel Allende's in the Midst of Winter. *Sada Al-Jamia Journal for Humanities*, 3(1), 124-139.
  11. Aisha Mohamed Ahmed (2025). Exploring Code-Mixing and Code-Switching in Bilingual English Speakers. *Comprehensive Journal of Humanities and Educational Studies*, 1(1), 01-10
  12. Aisha Mohamed Ahmed (2025). Comparative Analysis of Formative Assessment Strategies and Their Impact on Student Performance and Engagement. *Journal of Scientific and Human Dimensions*, 1(1), 01-09.
  13. Mohamed Belrzaeg (2025). Learning Analytics in E-Learning: Predicting Student Success through Data. *Journal of Scientific and Human Dimensions*, 1(1), 10-18.
  14. Karima Alfitory Ahmed. (2024). Prevalence and Awareness of Aphthous Ulcers Among Libyan University Students . *Libyan Journal of Contemporary Academic Studies*, 2(2), 35-44.
  15. Aisha M. Ahmed. (2025). Examining the Effectiveness of Distance Education: Challenges, Opportunities, and the Future of Learning. *Libyan Journal of Educational Research and E-Learning (LJERE)*, 1(1), 23-30.
  16. Asma Issa Karwad. (2025). Challenges Students Encounter When Translating Idiomatic Expression. A Case Study of Faculty of Education Students. *Journal of Libyan Academy Bani Walid*, 1(2), 229–242

**Disclaimer/Publisher's Note:** The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of **SAJH** and/or the editor(s). **SAJH** and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.