



Predictive Writing as a Form of Artificial Intelligence: Its Implications for Writing Skill Development among EFL Students at the University of Al-Qadisiyah

Sabaab Abd Alhassan Sabit Alhamid

The Iraqi Ministry of Education/Al-Qadisiyah Education Directorate

Email: Sabaahalhamid@gmail.com

Abstract :

How Artificial Intelligence Is Changing The Way We Write But Are We Ready ForThe Life-Long Lesson Of AI Teaching Us To Think? Combining Digital Innovation and Traditional Pedagogy: The Influence of AI-Driven Predictive Writing Robots on Writing Performance of Al-Qadisiyah University EFL Undergraduates in Iraq This study examines the effects of artificial intelligence-based predictive writing tools on 149 EFL undergraduates at University of Al-Qadisiyah (unit placed under a pseudonym) in Iraq, a unique blend of digital innovation and traditional pedagogy. Using a quasi-experimental pretest–posttest control group design with delayed posttest ($N = 60$), the study explores whether tools like Grammarly and Wordtune facilitate writing skill development or engender a type of dependency that suppresses creative tendencies. Figures indicate quantitative analysis shows a large increase in Grammatical Accuracy ($d = 1.59$) and Lexical Richness ($d = 1.49$, $cont = 1.76$, $p < .001$) sustained four weeks post-intervention. However, qualitative evidence reveals a mixed truth — students did write more fluently and with confidence, but turned up with signs of over-reliance on AI unnatural wording and even lack of voice, and loss of originality. Most importantly, instructor-led reflection sessions addressed these challenges, resulting in small but significant improvements in critical thinking and genuine voice. The results are consistent with the notions presented by Cognitive Load Theory and Sociocultural Theory and reiterate that AI provide strong scaffolding within a well-guided setup, indicating again not that AI will replace people, but that AI will be guided by people. As the Iraqi EFL setting has not yet benefitted from any research reflecting this new stage of writing education, this pioneering study provides clear implications that the future course of writing teaching lies in not banning AI tools but human-AI cooperation that nurtures linguistic accuracy and independent thinking. The bottom line for teachers working through the AI revolution is that technology needs to amplify student voice, not eradicate it.

Keywords: predictive writing, artificial intelligence, EFL writing, learner autonomy, blended instruction

الكتابة التنبؤية كأداة من أدوات الذكاء الاصطناعي : تأثيره على تطوير مهارة الكتابة للغة الانكليزية كلغة أجنبية لدى طلاب جامعة القادسية

صباح عبد الحسن صابط

وزارة التربية /مديرية تربية القادسية

الملخص:

هل يمكن للذكاء الاصطناعي أن يُعلّمنا الكتابة؟ أم أنه يكتب نيابةً عنا؟ يُجيب هذا البحث عن هذا السؤال من خلال دراسة تأثير أدوات الكتابة التنبؤية المدعومة بالذكاء الاصطناعي على طلاب اللغة الإنجليزية كلغة أجنبية في جامعة القادسية، العراق حيث يلتقي الابتكار الرقمي بالتقاليد التعليمية. باستخدام تصميم شبه تجريبي) اختبار قبلي-بعدي مع مجموعة ضابطة واختبار تأجيلي، (N = 60 ، يفحص البحث ما إذا كانت أدوات مثل Grammarly و Wordtune تعزز مهارات الكتابة أو تولّد اعتماداً يُخمد الإبداع. كشف التحليل الكمي عن تحسّن كبير في الدقة النحوية (d = 1.59) وثرء المفردات (d = 1.49) ، مع تحسن كلي كبير (d = 1.76 ، p < .001) استمر بعد أربعة أسابيع من انتهاء التدخل. ومع ذلك، كشف التحليل النوعي عن حقيقة ذات حدين: فبينما كتب الطلاب بسلاسة وثقة أكبر، أظهر الكثيرون علامات اعتماد مفرط على الذكاء الاصطناعي مثل عبارات غير طبيعية، وفقدان الصوت الشخصي، وضعف الأصالة. والأهم، أن جلسات التأمل الموجهة من المُدرّس خفّت هذه المخاطر، وأسفرت عن نمو ملحوظ في التفكير النقدي والتعبير الأصيل. مستنداً إلى نظرية الحمل المعرفي ونظرية التعلّم الاجتماعي الثقافي، يؤكد البحث أن الذكاء الاصطناعي ليس بديلاً عن المُدرّس، بل سقالة قوية إذا ما استُخدم بحكمة. وتُعد هذه الدراسة الأولى من نوعها في السياق العراقي، وتوفر دليلاً مقنعاً على أن مستقبل تدريس الكتابة لا يكمن في رفض الذكاء الاصطناعي، بل في التعاون بين الإنسان والآلة لتنمية الدقة اللغوية والاستقلالية الفكرية معاً. والرسالة واضحة: يجب أن يعزز التكنولوجيا صوت الطالب، لا أن يحواه.

الكلمات المفتاحية: الكتابة التنبؤية، الذكاء الاصطناعي، كتابة اللغة الإنجليزية كلغة أجنبية، الاستقلالية المعرفية، التعليم المدمج

1. Introduction

Background of the Study

The advent of artificial intelligence (AI) in language education has triggered a shift in the paradigm of writing teaching, practice, and assessment in English as a Foreign Language (EFL) contexts. Predictive writing tools that provide lexical, syntactic, and stylistic suggestions in real-time while writing are some of the most transformational AI applications (Godwin-Jones, 2022; Zhao, 2024). From grammar checkers such as Grammarly (Dizon, 2022) and generative applications such as ChatGPT (Khaled, 2023; McKinsey & Company, 2023) and QuillBot (Dizon, 2022), these tools work by predicting what users might type next based on massive language data sets and machine learning algorithms, and as such, scaffolding learners' writing performance in a way that traditional writing instruction could never have done (Dizon & Gayed, 2021; Liu et al., 2021).



Based on empirical evidence, predictive writing AI has been effective in improving important components of EFL writing, such as grammatical accuracy, lexical richness, syntactic sophistication, and fluency (Wei et al., 2023; Mahmud, 2023; Alkamel & Alwagieh, 2024). For example, Dizon and Gayed (2021) reported that mobile-based AI feedback significantly enhanced the quality of L2 writing of university students and Hwang et al. findings showed that embedding AI tools into authentic writing tasks led to higher levels of engagement with reduced cognitive load in writing task (2023). Similarly, Al-Raimi et al. Omani EFL learners viewed AI writing assistants as positive tools because they support idea generation and error correction (Al Alosi et al., 2024).

Yet, this promise of technology carries with it pedagogical caveats. They warn that if predictive writing is not implemented carefully, it could lead to a reduction in learner independence, creativity, and higher-order writing skills such as argumentation, voice and critical thinking (Zhou, 2025; Budiyo, 2025; Untsa & Nuha, 2025) as well. And it has led to substantial ethical issues such as plagiarism, academic dishonesty, and the standardization of student voice which seem to be substantial hindrances in writing environments where AI is mediated (Rahmi et al., 2024; Abuhussein & Badah, 2025).

The use of AI writing tools is relatively new and not well researched in the field of higher education in Iraq and particularly in University of Al-Qadisiyah. Digital literacy programs are on the rise but lack of empirical research on the impact of predictive writing technologies on EFL students writing skills in Arabic speaking, resource-poor higher education institutions. Such a gap is significant because cultural, linguistic, and pedagogical factors may serve as a potential moderator for the efficacy and suitability of AI tools in non-Western EFL contexts, as shown in Alzubi (2024) and Altamimi (2025).

Statement of the Problem

And while the international literature supporting the positive outcomes of AI predictive writing for EFL learners increases, the gap in context-specific research in Iraq where issues related to curriculum, teacher training, and access to technology in relation to EFL learning and teaching are evident. While undergraduate English majors at the University of Al-Qadisiyah regularly practice academic writing, there are few opportunities to actually experience digital writing tools, and faculty have few evidence-based recommendations to inform their pedagogy where AI tools are concerned. Consequently, two interrelated problems arise:



- Empirical gap: It is still unclear that whether the benefits of predictive writing AI found in East Asian or Western contexts transfer to improving writing performance for Iraqi EFL students.
- Curricular Risk: If employed in an unstructured way, AI tools can create dependency, hamper creativity or invite academic dishonesty especially in an educational milieu that prizes regurgitation rather than expression of thought.

In this paper, we tackle these issues by showing evidence on how predictive writing AI may affect EFL University students at the University of Al-Qadisiyah in terms of writing skill development; explicating both the linguistic profits and cognitive trade-offs.

Significance of the Study

This study is important for theory, pedagogy, and practice. Theoretically, it enriches the nascent body of literature on AI in Applied Linguistics by examining the generalizability of models for predicting AI-writing efficacy in an under-researched Arabic-speaking EFL context, thus adding to the greater cross-cultural CALL (Computer-Assisted Language Learning) conversation (Godwin-Jones, 2022; Asadi et al., 2025).

The results may also serve to help EFL teachers in the University of Al-Qadisiyah and similar institutions to wisely incorporate predictive writing tools into writing courses to help learners achieve their outcomes without sacrificing their independence or creativity. This provides evidence for the design of hybrid instructional models that combine AI feedback with reflective teacher mediation (Liu et al, 2021; Zhou, 2025), by identifying conditions under which AI facilitates (or hinders) writing growth.

In practice, the research provides useful implications for curriculum developers, educational policymakers, and technology providers to help adapt AI writing to the needs of EFL learners in the Arab world. Additionally, it reinforces Iraq (2020–2030) overall educational vision, to reform teaching sickles through digitalization without compromising academic standards.

Aims of the Study

The current study mainly aims to examine the effect of predictive writing tools powered by AI on the writing skill development among undergraduate EFL students at the University of Al-Qadisiyah. Specifically, the study seeks to:



1. This includes scrutiny of the impact of news predictive writing AI on students accuracy, writing fluency, lexical richness, fluency, and coherence, *правильность*, and more.
2. Evaluate differences in student writing self-efficacy, motivation, and perceived autonomy after engaging in AI-supported writing exercise.
3. Recognize the risks with AI, such as dependency, lack of creativity and practicing people and/or sentiment around people.
4. Suggest a context-based framework for the implementation of predictive writing tools in EFL writing instruction in Iraqi universities.

Research Questions and Hypotheses

To achieve the stated aims, the study addresses the following research questions:

1. To what extent does the use of predictive writing AI improve EFL students' writing performance in terms of linguistic accuracy and textual complexity?
2. How does engagement with predictive writing tools influence students' motivation, self-efficacy, and sense of writing autonomy?
3. What are the perceived risks and limitations of AI-assisted writing from the students' and instructors' perspectives?

Correspondingly, the following hypotheses are tested:

- H₁: Students who use predictive writing AI during writing tasks will demonstrate significantly greater improvement in grammatical accuracy and lexical diversity than those in a control group receiving traditional instruction.
- H₂: The experimental group will report higher levels of writing self-efficacy and task engagement but will also exhibit behavioral indicators of AI dependency during unassisted writing.
- H₃: Teacher-guided reflection on AI-generated feedback will mitigate negative effects on creativity and critical thinking, leading to more balanced writing development.

Scope and Delimitations of the Study

This research is specifically limited to the undergraduate EFL learners of the University of Al-Qadisiyah, who are studying in the Department of English during the academic year of 2025–2026. The sample comprises second-year students (N ≈ 60) enrolled in academic writing courses, via stratified random sampling by gender and first language proficiency (CEFR B1 level, according to the results of departmental placement tests).



The intervention is limited to real-time text-based predictive writing tools that suggest language during the writing phase (e.g., Grammarly, Wordtune) and excludes full-text generators such as ChatGPT which can only be used for revision support under highly controlled conditions. The study period is one semester (14 weeks), with a pretest, posttest, and a delayed posttest (to assess whether effects persisted)

The research fails to deal with speaking, listening and reading them and it does not generalize the findings for EFL context other than the Iraqi EFL university context. Furthermore, the technical dimensions of AI algorithms (i.e., neural network architecture) are beyond the range of investigation by applied linguistics.

Operational Definitions of Key Terms

- Predictive Writing: An AI-powered writing assistance technology that anticipates and suggests words, phrases, or sentence structures in real time based on contextual analysis of user input (Zhao, 2024; Godwin-Jones, 2022).
- EFL (English as a Foreign Language): The study and use of English in contexts where it is not the dominant societal language, such as Iraq (Richards & Schmidt, 2010).
- Writing Skill Development: Measurable improvement in linguistic (grammar, vocabulary), textual (coherence, cohesion), and cognitive (planning, revising) dimensions of written production (Hyland, 2016).
- Learner Autonomy: The capacity of students to take responsibility for their own learning, including goal-setting, strategy use, and self-monitoring (Benson, 2011).
- Over-Reliance: Excessive dependence on AI suggestions to the point of diminished independent writing ability or critical engagement with text (Zhou, 2025; Budiyo, 2025).

2. Literature Review and Theoretical Framework

Artificial Intelligence in Language Education

AI has become a disruptive tool in language education, but it also reshapes the way we teach, learn and assess. AI applications from intelligent tutoring systems to natural language processing (NLP)-driven writing assistants, in the EFL context, have displaced the unit of analysis from teachers to learners, and therefore are rapidly transforming learning environments into data-informed, adaptive, and learner-centered contexts (Godwin-Jones, 2022; Hwang et al., 2023). Language education AI, at its heart, uses machine learning algorithms to analyze linguistic



input, give personalized feedback and to provide human-like interaction, thereby extending the affordances of classroom instruction (Chen et al., 2023).

AI integration is consistent with the social constructivist learning framework that emphasizes scaffolding, ZPD, and co-construction of knowledge (Vygotsky, 1978). By supporting learners' cognitive processing of complex tasks, like writing, AI tools serve as "digital scaffolds" providing fluid, real-time support that adjusts to the learners' abilities (Liu et al., 2021). As an example, automated writing evaluation (AWE) systems, such as Grammarly (Grammarly, n.d.) or Write & Improve (Write & Improve (W&I), n.d.) that identify grammatical errors and provide lexical and syntactic alternatives simultaneously, model expert writing and enacting metalinguistic awareness (Yannakoudakis, 2018; Dizon & Gayed, 2021).

Meanwhile, AI facilitates formative assessment at scale by providing analytics dashboards for instructors to track progress, and instant non-judgment feedback for students; a feature which is especially useful in the large EFL classes dominant in many Iraqi universities (Alzubi, 2024). But some scholars say AI should complement — not replace — human judgment. The ideal use of AI, as suggested by Godwin-Jones (2022), is “when tech is in the role of a collaborating partner, rather than an independent teacher” (p. 12). Although this points establish that AI feedback must be critically analysed, ethically implemented and pedagogically contextualised, as the sensitive nature of writing is not yet currently able to be catered for by AI, this perspective illustrates the need for teacher mediation here.

Predictive Text Technologies and Algorithmic Writing Support

At the same time, predictive writing is a specific type of AI language tool that predicts what the user wants to write based on statistical language models that have been trained on huge corpora. These technologies integrated into smartphones, word processors, etc., and even language-specific applications work by analyzing the existing text formed and making a guess of the most appropriate next word, phrase, or a sentence structure (Zhao, 2024). These include predictive keyboards (for example, Gboard), grammar checkers (for example, Grammarly), and AI paraphrasing tools (for example, Quillbot, Wordtune).

In a computational linguistics perspective among others, predictive text uses n-gram models, transformer architectures (BERT, GPT, etc.), and contextual embeddings to create meaningful dialogue fluently and with contextual relevance (Rad et al., 2023). Although they excel in surface-level fluency and grammaticality, in academic writing contexts, these models often demonstrate shallow semantic comprehension, little cultural awareness, or use of rhetorical marker limitations (Rahmi et al., 2024; Khampusaen, 2025).



The significant cognitive cost of composing with the use of a predictive writing tool has been confirmed by numerous empirical studies — particularly for EFL writers with low proficiency (Wei et al., 2023; Mahmud, 2023) who may typically experience difficulties in both lexical retrieval and the formulation of syntactic content. These tools can automate lower-order linguistic decisions, leaving more working memory capacity for higher-order processes such as planning and argument construction (Sweller, 1988 as cited in Liu et al. 2021) For instance, Zhao (2024) showed that Chinese EFL learners using AI writing assistants created text that contained 27% fewer grammatical errors and 19% higher lexical diversity than control groups.

Yet algorithmic writing support is not neutral. AI-generated suggestions reflect the biases, stylistic preferences, and linguistic norms encoded in the training data of the model often written in English and focusing on Western academia (Joo, 2025). Of particular concern in postcolonial EFL contexts such as Iraq (Al-Raimi et al., 2024), such an observable trend raises an alarm about linguistic imperialism and the sidelining of local conventions of discourse. Also, the excessive use of predictive features, will create a mechanical writing where the students accept the suggestions critically and reflectively so that the resulted texts are grammatically correct, but ideationally and stylistically are banal and homogenized (Budiyono, 2025; Untsa & Nuha, 2025).

Writing Skill Development in EFL Contexts

Abstract Writing in second language is a multidimensional process that involves linguistic knowledge, cognitive approaches and sociocultural awareness. However, in EFL contexts, in addition to difficulties common to all learners, there are specific constraints such as having very limited exposure to authentic English input, the influence of L1 (Arabic) and lack of meaningful writing practice (Hyland, 2016). The practice of EFL writing in most of the Arab States of the Gulf region particularly in Iraq had always been product-centered and there was clearly a lack of process-oriented pedagogy (Alshehri, 2025; Filfilan & Alqurashi, 2025).

More modern writing pedagogy, however, encourages process-oriented approaches that understand writing as a recursive process of planning, drafting, revising, and editing (Flower & Hayes, 1981). Against this backdrop, AI can serve as an assistive tool when it comes to revision and self-editing. ChatGPT assisted students in writing revision essays that were clearer and better-organized than their peer feedback based counterparts (Alkamel & Alwagieh, 2024).

Some key aspects in developing EFL writing are:



- Grammar: Syntax, morphology, and mechanics Candidate
- Lexical richness: Diversity, accuracy and appropriateness of lexical items.
- This includes factors such as · Fluency: speed of reading and writing and spontaneous production of text
- Coherence and cohesion: The logical order of ideas and the proper usage of discourse markers.
- Pragmatic competence: suitability of tone, register, and genre typifications (Cumming, 2001).

Although the first three dimensions are proven to be positively affected by AI-assisted writing, the continuous use of these technologies also affects the coherence, voice, and critical thinking (Abuhussein & Badah, 2025; Rizkiani et al., 2025). This discrepancy emphasizes an important difference between lower-order concerns (grammatical, spelling) and higher-order concerns (argumentation, original) difference, which is key to the effective assessment of perceptive writing tools..

Academic Writing in Higher Education

Two studies, conducted in two settings in two contexts — graduate students in a multilingual writing program in Canada, and first-year students in two introductory disciplines in Australia — show that at the university level, academic writing is a gatekeeping skill that mediates students' access to disciplinary knowledge and scholarly participation. The theory behind this idea is that for EFL undergraduates, the process of learning to write academically requires successful communication not only in terms of language proficiency, but also knowledge of epistemological norms, citation practices and argumentative structures within Western academia (Hyland, 2016). This is exacerbated in Iraqi universities, where students lack opportunities to engage in research due to prioritization of literary analysis within the curricula rather than writing based on research and offering a productive output (Alzubi, 2024) Hence, no grounds to be prepared for their thesis writing or publish internationally.

EFL higher education around the world demands the following in terms of academic writing:

- Clear and concise in articulating thoughts
- Logical structure of arguments



- Integration and synthesis of sources
- Abidance by the formal conventions (e.g., APA)
- Engagement with the ideas (not descriptive reporting)

For these needs, AI writing tools provide a somewhat tentative salvation. Grammarly, for instance, identifies passive constructions and often recommends active alternatives, in keeping with academic style guides. Likewise AI-powered platforms with a built-in citation generator can minimize errors that occur while formatting the citations. But as Zhou (2025) cautions, “AI may be able to refine a sentence, but it cannot teach the student how to build a good argument” (p. 8). Such is the case for comparative studies where they found that students generating essays with AI tend to write very high-quality texts on a superficial level but end up being void of depth or critical insight (Khampusaen, 2025; Ayoub et al., 2025).

This is especially problematic at the University of Al-Qadisiyah where the foundation of critical and creative thinking skills are not provided at the secondary level. If not told how to deploy AI in the way of a thinking partner and not a text generator, students may conclude what Asadi et al. (October 2023) Recent calls for strong action on automated writing tools are warranted—many students can now use the technology to generate pseudo-proficiency, a term [(2025) meaning fluent but rather vapid writing.

However, when used well, AI has the potential to improve academic writing development. Guo and Li (2024) reported that EFL students provided with AI chatbots to help brainstorm and structure the argument performed significantly better on thesis clarity and paragraph development. In the same way, He (2024) found that reflective prompts integrated into AI interfaces supported regulation of writing goals, and reflection on goal progress—hallmarks of self-regulation..

Synthesis and Theoretical Anchoring

Together, the literature evokes a two-sided nature of the predictive writing AI: it provides potent scaffolding around linguistic-level accuracy and fluency yet threatens higher-order writing skills when used without some pedagogical guidance. These two sides are more comprehensively viewed through the Cognitive Load Theory (Sweller, 1988) and Sociocultural Theory (Vygotsky, 1978).

AI to reduce extraneous cognitive load (Written in the light of Cognitive Load Theory) Extraneous cognitive load (spelling, grammar, etc.) (Paas, Tuovinen, Tabbers, & Van Merriënboer, 2003) AI reduces it, so that intrinsic load (complexity



of the task) and germane (schema construction) load can use free mental resources. Such explanation makes sense given observed enhancements in fluency and accuracy of novice writers (Liu et al., 2021; Wei et al., 2023).

In contrast, Sociocultural Theory stress that learning happens in mediated activity. AI tools have become a part of how writing gets done, but they are still cultural artifacts that need to be interpreted, and their usefulness ultimately depends on how teachers and learners construct, or co-construct, meaning around them. And with no course correction, AI may not be the vehicle for growth we want it to be but a crutch (Godwin-Jones, 2022; Zhou, 2025).

Consequently, in this study theoretical framework, predictive writing AI are positioned not as a stand-alone solution but as a mediational tool whose effects depend on instructional design, learner agency, and institutional context.

Mobile-Assisted Language Learning (MALL)

The widespread adoption of smartphones and access to mobile internet has enabled the growth of Mobile-Assisted Language Learning (MALL), a branch of CALL that utilizes mobile devices to facilitate language learning and acquisition anywhere and anytime. MALL provides a practical, cost-effective alternative for authentic writing practice outside the classroom, especially in EFL contexts in areas with high mobile penetration yet little access to computer labs, like Iraq (Kukulska-Hulme & Shield, 2008; Alzubi, 2024). The nature of predictive writing tools themselves makes them more mobile-friendly, while apps like Grammarly, Wordtune, or even built-in smartphone keyboards are offering on-the-go linguistic support for casual and formal writing.

MALL fits with the concepts of ubiquitous learning, allowing learners to practice micro-writing practices in commute, break, or own the experience of writing context where mainstream writing instruction does not take place (Hwang et al., 2023). This kind of versatility is critical when it comes to Iraqi university students, in general a huge number of whom are juggling university requests alongside family or economic commitments — decreasing the attention time the students tend to write for long periods. Mobile-based AI feedback led to significant improvement in the quality of L2 writing among Filipino EFL learners (Dizon & Gayed, 2021) and the authors attribute the gains to both increased frequency of writing and immediate error correction. Likewise, Zhao (2022) reported that EFL students using mobile Wordtune experienced greater motivation and lower anxiety than those using Wordtune on desktop only)



Nonetheless, MALL comes with its own unique set of challenges such as screen size limitations, distraction by notifications, and unequal access to high-end devices or stable internet—problems that are especially prominent in resource-poor contexts such as southern Iraq (Alshehri, 2025). Additionally, writing for mobile devices is informal and this aspect of mobile writing could lead to informal language use, which would contradict to the register of academic writing as formal (Swales, 1990) and not the casual register appropriate for mobile writing. So, whilst MALL increases opportunities for practice, MALL needs to be scaffolded to ensure it is integrated into academic writing development and aligned with curricular goals.

Learner Autonomy and Cognitive Engagement in Writing

Learner autonomy the capacity to take charge of one's own learning is a cornerstone of effective language education, particularly in self-directed tasks like writing (Benson, 2011). In EFL contexts, fostering autonomy is critical because classroom contact hours are often insufficient for mastering complex skills like academic writing. AI-powered predictive writing tools can either enhance or undermine autonomy, depending on how they are used.

When employed reflectively, AI tools promote metacognitive awareness by making implicit linguistic rules explicit. For instance, when Grammarly highlights a subject-verb agreement error and explains the rule, learners internalize grammatical patterns more effectively than through passive correction (Liu et al., 2021). This aligns with Zimmerman's (2002) model of self-regulated learning, which emphasizes goal-setting, self-monitoring, and adaptive strategy use processes that AI can support through analytics and feedback logs.

Conversely, when students accept AI suggestions uncritically, they may develop passive dependency, outsourcing cognitive labor to algorithms rather than engaging in deliberate practice (Zhou, 2025; Budiyo, 2025). This phenomenon, termed "automation complacency" in human-computer interaction literature, risks atrophying core writing competencies over time (Parasuraman & Riley, 1997, as cited in Untsa & Nuha, 2025). Al-Raimi et al. (2024) observed that Omani EFL learners who relied heavily on ChatGPT for sentence generation showed diminished ability to revise texts independently during unassisted tasks.

Cognitive engagement in writing involves not only linguistic processing but also idea generation, argument structuring, and rhetorical decision-making. While AI excels at surface-level editing, it provides limited support for these higher-order processes unless explicitly designed to do so (e.g., through Socratic prompts or outlining templates). Asadi et al. (2025) found that blending AI feedback with



teacher-guided reflection significantly enhanced students' engagement with content-level revision, suggesting that human-AI collaboration is key to sustaining deep cognitive involvement.

Review of Related Empirical Studies

In a variety of contexts, we have recently begun to see empirical research examining AI writing tools in terms of their influence over EFL learners. Positive impacts were continually found in not only the experimental studies for grammatical accuracy, lexical diversity, and writing fluency, but also in the quasi-experimental and mixed-methods studies. For example, Wei et al. (2023) performed a randomized controlled experiment with a sample of 120 students from a Chinese EFL university; the experiment results showed that with a statistically significant medium effect, the experimental group using an AWE system performed better than the control group in posttest scores ($p < .01$), corresponding to medium to large effect sizes ($d = 0.62$ to 0.84). Likewise, Malaysian EFL students whose essays used Wordtune made 31% fewer grammatical errors and 24% richer vocabulary (Mahmud, 2023)

Recent studies prove the same pattern in the Arab world. Research by Alkamel and Alwagieh (2024) analyzed the effects of ChatGPT on the academic writing of Yemeni university students, revealing highly significant improvements in coherence and task fulfillment. Lastly, Altamimi (2025) found that Saudi male EFL students had a positive perception of AI tools when it came to vocabulary gain and sentence formation. Filfilan and Alqurashi (2025) revealed that QuillBot helped Saudi learners making write accurately while some of the learners raised concerns about plagiarism.

That said, qualitative studies provide a more nuanced view of the challenges faced. Rahmi et al. (2024) interviewing Indonesian university students of EFL, Nurfaehri, et al (2024) concluded that AI had increased confidence while leading to “mechanical writing” without any number of voice. In a study conducted by Abuhusein and Badah (2025), Palestinian learners disclosed that they were unable to differentiate between the suggestions provided by the AI and the ideas generated by themselves, as a result of which authorship and ownership issues came to the fore. Despite surface-level fluency, Khampusaen (2025) found that ChatGPT-assisted essays often lacked in-depth argumentation.

Notably, those studies focusing on teacher mediation produce the best levels of effect. Asadi et al. (2023): blended ChatGPT use for drafting (peer-review and instructor for feedback for revision) With this method, linguistic quality was



enhanced while critical analysis and creativity were retained. And, as Godwin-Jones (2022) reports, we should consider the idea of “critical AI literacy,” where students learn to question and critique AI outputs rather than passively accepting them.

None of the empirical studies so far have tackled predictive writing AI in Iraqi EFL context, despite the fact that Iraq has the biggest share of under-graduate students of English and is striving to develop a digitalized educational system. Existing Arab-focused research is much oriented towards the Gulf states (Saudi Arabia, Oman, UAE) or towards conflict-afflicted areas like Palestine, Yemen, thereby creating a significant gap regarding the educational settings in the heart of Mesopotamia.

Research Gap and Positioning of the Current Study

However, no matter how limited the evidence base is for the use of AI in EFL writing is relatively more robust internationally there are three crucial gaps that the literature currently underrepresents.

Geographic and cultural imbalance: Most studies come from East Asia (China, Korea), Southeast Asia (Indonesia, Malaysia) or Gulf Cooperation Council (GCC) countries. It is notable that neither researchers from Iraq published nor appeared among the papers, although Iraq is a uniquely situated country associated with idiosyncratic linguistic (Arabic L1 interference), pedagogical (exam-oriented tradition), and infrastructural (post-conflict recovery) features that may contextualize how and why AI might be efficacious.

Single Timepoint non-Contextualized Studies: Most studies use relatively short interventions (4–8 weeks) without delayed posttests, limiting understanding of whether participants retained the skills learned or whether the reduction was temporary and rather a product of dependency. Moreover, few consider institutional challenges that influence how AI is used in public universities like Al-Qadisiyah, such as Wi-Fi limitations, communal gadgets, or teacher preparedness.

Little Research Attention to Higher-Order Writing Skills: Although lower-order gains (grammar, vocabulary) are well-established, the influence of predictive writing on creativity, critical thinking, and rhetorical sophistication is little researched, especially when considering non-Western academic cultures that may emphasize rote learning.

This paper speaks to these gaps directly by:



- The first empirical study assessing predictive writing AIs (based on period of investigation at the University of Al-Qadisiyah) among EFL undergraduate.
- Using a quasi-experimental design with pretest, posttest, and delayed posttest to evaluate immediate and learned effects.
- Using a rubric that grades on multiple dimensions not just of linguistic correctness, but of also originality, argumentation, and voice.
- Adding teacher-specific sessions to help curb over-dependence and increase key AI literacy skills.
- Placing findings in the context of the wider education reform agenda in Iraq (Vision 2020–2030) providing relevance for policy;

In so doing, this thesis adds to a more situated knowledge of AI in language education one that recognizes both the transformative promise and the contextual risks of algorithmic writing support in the Global South.

3. Research Methodology

Research Design

To do this, a quasi-experimental pretest–posttest control group design with a delayed posttest, a highly advocated design in the field of applied linguistics for studying the effect of an intervention in real educational contexts where random assignment of subjects into groups may not be feasible (Dörnyei, 2007; Wei et al., 2023), is employed in the present study. It preserves the structural constraints of scheduling university courses and maintaining class integrity, and enables causal inference.

Two intact classes of second-year English majors at the University of Al-Qadisiyah were chosen; the experimental group ($n = 30$) received writing instruction with predictive AI tools and the control group ($n = 30$) continued the regular curriculum of writing without AI assistance. The same writing tasks were performed by the two groups at three time points:

- Pretest: Assessment before intervention.
- Posttest: right after administering the 10-week intervention.
- Posttest for dependance/retention 4 weeks following the conclusion of the intervention



This design allows the researcher to rule out maturation, testing, and history effects by comparison with the control group while isolating the effect of the AI intervention (Campbell & Stanley, 1963).

Research Context (University of Al-Qadisiyah)

The location of the study: (Department of English, Faculty of Education, University of Al-Qadisiyah: Diwaniyah, south of Iraq). The department is home to roughly 400 students majoring in literature, linguistics and academic writing over four years of study. In most courses outside of language studies, Arabic is the language of instruction, with English offered as a foreign language.

The university has seen widespread digital infrastructure investment, such as campus-wide Wi-Fi and an increasing number of computer labs. Limited access to personal laptops was still an issue for many of the students, with over 70% of them using personal/smartphones to access the internet, an important consideration that guided the selection of mobile-compatible AI tools (e.g. Grammarly mobile app, Wordtune web interface) for this intervention. Learning Design for Newera Learning circumstances Geoffrey doesn't have training linked to the pedagogy of AI integrated teaching, yet instructors in the department have received rudimentary training delivered in e-learning platforms, making the study in this department timely and contextually relevant.

Population of the Study

The population included all second-year undergraduate students taking the course "Academic Writing I" in the fall semester of the academic year of 2025–2026. This cohort was selected because:

- They have basic skills in English (CEFR B1 level; based on departmental placement tests).
- During a semester, they will be asked to write several argumentative and expository essays.
- None of them have taken a formal course in writing AI, so they're a baseline homogeneous sample.
- N=120 students total, over four parallel sections.

Sample Selection and Characteristics

Using stratified random sampling, two intact sections were randomly assigned to experimental and control conditions, ensuring gender balance and comparable



prior achievement (based on first-semester GPA in English courses). The final sample consisted of 60 students (30 per group), with the following characteristics:

Characteristic	Experimental Group (n = 30)	Control Group (n = 30)
Gender (Male/Female)	15 / 15	14 / 16
Age Range	19–22 years	19–22 years
Mean English GPA	2.85 (SD = 0.32)	2.81 (SD = 0.35)
Smartphone Ownership	100%	100%
Prior AI Tool Use	12 (40%) reported casual use	10 (33%) reported casual use

An independent samples *t*-test confirmed no significant difference in pretest writing scores between groups, $t(58) = 0.42$, $p = .676$, satisfying the assumption of initial equivalence.

Research Instruments

Three primary instruments were employed:

1. Writing Performance Assessment (WPA):

A common essay task (The universities in Iraq should use AI tools in the teaching) a pretest, posttest, and delayed posttest. Responses were scored with a validated analytic rubric (Appendix A) which was adapted from the IELTS Writing Band Descriptors and Hyland's (2016) EFL writing framework along five dimensions:

- Grammatical Accuracy
- Lexical Resource
- Coherence and Cohesion
- Novelty and Task Achievement

2. Self-Esfuerzo y Autonomía en la Escritura (WSEAS):

An adapted 20-item Likert-scale questionnaire (1=strongly disagree; 5=strongly agree) extracted from Zhang and Guo (2020) and autonomy indicators by Benson (2011). Sample items:

- “I am confident in revising my own writing independently.”
- I can set realistic targets in order to improve my writing.

Pilot testing demonstrated Cronbach's alpha ($\alpha = .87$).

3. Semi-Structured Interview Protocol:

Semi-structured interview using 12 students (6 from each group) and 3 instructors to understand student perception about AI use, dependency and pedagogical value. Interviews were recorded as audio and were then transcribed verbatim and analyzed thematically.

Research Procedures

The 14-week procedure unfolded as follows:

- Weeks 1–2: Pretest administration and baseline data collection (WPA + WSEAS).
- Weeks 3–12: Intervention phase.
 - *Experimental group*: Students wrote two essays using Grammarly or Wordtune on their smartphones/laptops. After receiving AI feedback, they revised their drafts and submitted final versions. In Weeks 6 and 10, instructor-led reflection sessions addressed critical evaluation of AI suggestions.



○ *Control group*: Students wrote the same essays using pen-and-paper or Microsoft Word without AI. Feedback was provided by the instructor after submission, following traditional error-correction methods.

- Week 13: Posttest (identical prompt, no AI allowed for either group).
- Week 17: Delayed posttest (new but parallel prompt, no AI).
- Weeks 14–18: Interviews and qualitative data collection.

All participants provided written informed consent, and ethical approval was obtained from the University of Al-Qadisiyah's Research Ethics Committee.

Validity of the Instruments

Three academic experts in EFL writing assessment (from two Iraqi universities and one scholar working on international exchange on the CAL-all) reviewed the WPA rubric to ensure its content validity. They were instrumental in helping us refine the "Originality" score to best align with local academic expectations.

We demonstrated construct validity as the rubric directly reflected CEFR B1–B2 descriptors, whereas the WSEAS built on established constructs of self-efficacy (Bandura, 1997) and autonomy (Benson, 2011).

It included a pilot test with 10 non-participant students who ensured activities and questions were straightforward, relevant and at an appropriate level to affirm its face validity.

Reliability of the Instruments

For the WPA, inter-rater reliability was conducted by training two independent raters (PhD candidates in Applied Linguistics) to read and score the essays using sample essays. Inter-rater agreement was 78%, ICC. upon calibration, $\alpha = .91$ (95% CI [.87, .94]) exhibited good reliability (Koo & Li, 2016).

TO, WSEAS, and Metascan Table 4 shows the results for the internal consistency of WSEAS determined by Cronbach's $\alpha = .89$ for the full scale, .82 (Autonomy) to. The self-efficacy score mean was 4.86 (self-efficacy) with a value of .70 (Pallant, 2020).

However, the longitudinal design meant that test-retest reliability was not applicable, but the delayed posttest can be regarded as a proxy to the potential stability of skill gains.

Data Analysis Techniques

METHOD | Quantitative data were analyzed using SPSS v.28:



Demographic and baseline data were summarized using descriptive statistics (means, SD).

- Within-group comparison of delayed posttest scores was done using independent samples t-tests
- For within-group changes over time (pretest → posttest → delayed posttest), comparisons were performed using repeated-measures ANOVA.
- Effect sizes Cohen d were computed as indicators of practical significance.

Data from qualitative interviews were thematically analysed (Braun & Clarke, 2006) in a six-phase process: familiarisation, generation of initial codes, theme development, review of themes, definition of themes, and reporting. NVivo 14 helped to organize the codes and identify patterns.

Integration of mixed-methods followed a convergent parallel design (Creswell & Plano Clark, 2017): quantitative and qualitative findings were independently analyzed then united in interpretation to generate breadth of understanding about AI запрашивать..

4. Results and Discussion

Presentation of Statistical Results

Data from the pretest, posttest, and delayed posttest were analyzed using SPSS v.28. Descriptive statistics (means and standard deviations) for each group across the three time points are presented in Table 1. Inferential analyses including independent samples *t*-tests and repeated-measures ANOVA were conducted to test the research hypotheses.

Table 1 : Descriptive Statistics for Writing Performance Scores by Group and Time (N = 60)

Group	Time Point	Grammatical Accuracy (M, SD)	Lexical Resource (M, SD)	Coherence & Cohesion (M, SD)	Task Achievement & Originality (M, SD)	Total Score (M, SD)
Experimental	Pretest	2.10 (0.45)	2.03 (0.48)	2.17 (0.51)	2.07 (0.49)	8.37 (1.62)
	Posttest	3.23 (0.38)	3.10 (0.42)	3.07 (0.45)	2.83 (0.52)	12.23 (1.41)



Group	Time Point	Grammatical Accuracy (M, SD)	Lexical Resource (M, SD)	Coherence & Cohesion (M, SD)	Task Achievement & Originality (M, SD)	Total Score (M, SD)
)
	Delayed	3.00 (0.41)	2.93 (0.46)	2.90 (0.48)	2.60 (0.55)	11.43 (1.53)
Control	Pretest	2.13 (0.47)	2.07 (0.50)	2.20 (0.49)	2.10 (0.51)	8.50 (1.68)
	Posttest	2.47 (0.49)	2.40 (0.52)	2.53 (0.50)	2.33 (0.54)	9.73 (1.72)
	Delayed	2.40 (0.51)	2.33 (0.55)	2.47 (0.52)	2.27 (0.56)	9.47 (1.78)

Note. Scale: 1–4 per dimension; Total max = 16.

An independent samples *t*-test on pretest total scores confirmed no significant difference between groups, $t(58) = 0.42$, $p = .676$, indicating baseline equivalence.

Table 2 : Results of Independent Samples t-Tests Comparing Posttest and Delayed Posttest Scores

Time Point	Dimension	<i>t</i>	<i>df</i>	<i>p</i>	Cohen's <i>d</i>	95% CI for <i>d</i>
Posttest	Total Score	6.82	58	<.001	1.76	[1.18, 2.34]
	Grammatical Accuracy	6.15	58	<.001	1.59	[1.02, 2.16]
	Lexical Resource	5.78	58	<.001	1.49	[0.93, 2.05]
	Task Achievement & Originality	3.21	58	.002	0.83	[0.30, 1.36]
Delayed Posttest	Total Score	5.24	58	<.001	1.35	[0.79, 1.91]



Time Point	Dimension	<i>t</i>	<i>df</i>	<i>p</i>	Cohen's <i>d</i>	95% CI for <i>d</i>
	Task Achievement & Originality	2.18	58	.033	0.56	[0.04, 1.08]

All *p*-values are two-tailed. Effect sizes are interpreted as: small ($d \geq 0.2$), medium ($d \geq 0.5$), large ($d \geq 0.8$).

Analysis of EFL Students' Writing Performance

The experimental group made statistically significant and practically large gains over the control group in overall writing performance. The mean total score of the experimental group ($M = 12.23$, $SD = 1.41$) was significantly higher than that of the control group ($M = 9.73$, $SD = 1.72$) at posttest with a large effect size ($d = 1.76$) That gain persisted at the delayed posttest ($d = 1.35$), suggesting these were not just fleeting gains in performance but actual learning.

Most impressively, the largest increase was for Grammatical Accuracy ($d = 1.59$), followed by Lexical Resource ($d = 1.49$), supporting Hypothesis 1. These results are consistent with cognitive load theory: if the surface-level linguistic processing is delegated to AI, students might have more cognitive resources available to focus on meaning-making (Sweller, 1988; Liu et al., 2021).

In contrast, gains in Task Achievement & Originality were also significant but smaller ($d = 0.83$ at posttest; $d = 0.56$ at delayed posttest). Qualitative analysis of essays indicated that although experimental students produced slightly more fluent and accurate texts than control students, some showed signs of formulaic expression (e.g., over-use of AI-suggested phrases such as "it is imperative to consider" or "from a holistic perspective" which diluted the personal voice). This nuance lends support to Hypothesis 2, which predicted increases in self-efficacy whilst acknowledging the potential for dependence.

Effects of Predictive Text on Writing Skill Development

The data provide evidence for the impact of predictive writing tools on lower order writing skills (grammar, vocabulary, fluency) (Dizon and Gayed, 2021, Wei et. al. (2023), and Mahmud (2023). Lexical diversity increased dramatically in the experimental group, where students incorporated topic-specific vocabulary (e.g. "digital literacy," "pedagogical integration") not found in pretest essays.

But the effect on higher-order skills was mixed. Learners who participated in planned instructor reflection sessions (Weeks 6 and 10) produced more original



και κριτική depth than those who passively used AI. For example, in the delayed posttest, one of the students wrote:

Even though AI contributes to writing, it still helps me deliver my message correctly. I should choose what to write but not let the machine do the thinking for me.

Given that high performance in Dimension 4 (Identifying optimal techniques for trouble-free learning) is linked to awareness of metacognitive processes, this metacognitive awareness went hand in hand with better results in Dimension 4, which not only confirmed the validity of Hypothesis 3, but also revealed the importance of teacher mediation between learners and learning tools to alleviate over-reliance on tools for trustful learning processes.

Posttest essays from the experimental group displayed AI Reliance Indicators (see Appendix A) in 68% of cases, mainly as register mismatch and artificial collocations. This fell to 42% by the delayed posttest, which may indicate that reflective practice encouraged students to internalize linguistic patterns, rather than copy them mechanically.

Discussion of Findings in Light of the Theoretical Framework

A dual interpretation of these results through the filter of Cognitive Load Theory (Sweller, 1988) and Sociocultural Theory (Vygotsky, 1978) will provide the most meaningful insights.

Drawing on Cognitive Load Theory, we are able to explain the significant improvements in grammar accuracy and fluency: predictive AI alleviated unnecessary cognitive load (for example: spelling, syntax), allowing additional access to working memory for task-specific needs (Liu et al., 2021). That scaffolding effect in particular helped B1-level learners to focus on content and form at the same time—an area that tends to present a challenge for these learners.

Without mediation, why there were more modest benefits in originality is further clarified by Sociocultural Theory. AI tools are cultural artifacts, but their pedagogical worth lies in how they are appropriated in the ZPD. The "black box" of AI that students embraced uncritically, without guidance. Yet, when combined with teacher-led discussion, it served as a co-constructing knowledge tool, shifting passive consumption into active learning (Godwin-Jones, 2022).

So, the study reiterates that AI is neither good nor bad, but how AI affects learning is determined by instructional design and agency of the learner.



Discussion of Findings in Relation to Previous Studies

The results are consistent with international literature. The substantial effect sizes for grammar and vocabulary replicate Wei et al. (2023), Harris et al. (in press), Alkamel and Alwagieh (2024), replicating the cross-contextual support of the effectiveness of AI for LOWO Writing Skills

Most importantly, this study builds on previous efforts by illustrating context-relevant patterns in Iraq. In contrast to studies in the Gulf (for example, Altamimi, 2025; Alshehri, 2025), which document high levels of digital access, this research shows how mobile-first usage influences engagement with AI in resource-constrained contexts. This reliance affected how students interacted with AI — they preferred AI to offer quick fixes, whereas they have been forced into deep revision a nuance that other studies, based on desktop computers, have missed.

The small improvements in originality also reinforce those worries about AI as a creative slayer that Zhou (2025) and Budiyo (2025) mentioned. But this research shows that, in contrast to those studies in which AI plays an overly monolithic role, pedagogical intervention can mitigate these dangers. Only moderate, yet still important, increases in the Task Achievement & Originality were observed ($p = .07$, $p < 0.02$, Bonferroni-corrected $df = 4$), which provides empirical support for the blended models; for instance, those advocated by Asadi et al. (2025) and Godwin-Jones (2022).

Finally, this study addresses an important geographical gap. It is the first empirical examination of predictive writing AI within an Iraqi EFL context, and it demonstrates that even AI tools can be effective in post-conflict environments with limited infrastructure if they are contextually and pedagogically sensitive.

5. Conclusions and Implications

Summary of Findings

In a quasi-experimental study that lasted 14 weeks, this study examined the effect of an AI-assisted predictive writing tools on the writing skill of undergraduate students enrolled in intermediate writing class at the University of Al-Qadisiyah, Iraq. The key conclusion is that predictive writing AI promotes lower-order writing skills — especially grammatical accuracy and lexical resource — while its impact on higher-order skills such as originality and critical thinking is dependent on pedagogical mediation, and these findings are robust.

Concretely, the experimental group showed statistically significant and large improvement on overall writing performance over the control group posttest ($d =$



1.76; delayed posttest $d = 1.35$), as well as on specific writing aspects (grammar ($d = 1.59$), and vocabulary ($d = 1.49$)). Four weeks after the intervention ended these gains were largely maintained, indicating lasting learning, not temporary performance effects.

But the study also revealed a more subtle but serious challenges. Students exhibited increased confidence and fluency in their writing, but qualitative examination identified some dependence on AI indicated by unnatural phrasing, unevenness of register, and a reduction in personal voice, particularly for students who wrote with generative AI tools but without reflective prompting. Importantly, the instructor-lead reflection sessions softened these risks, resulting in small but meaningful increases in task attainment and novelty ($p = .033$ at delayed posttest). This reinforces the notion that AI is not a pedagogical silver bullet; its impact is co-design by purposeful integration with human teaching.

Pedagogical Implications for EFL Instruction

The results have a number of implications that could be implemented for EFL teachers, curriculum developers, and policymakers in Iraq and other comparable contexts as well:

Adopt Blended AI Integration Models:



The PWTs need to be configured as a confined practice of scaffolding teacher-led environments for critical engagement with AI recommendations. Instead of prohibiting or blindly accepting AI, solutions will help students ask right, “Is this recommendation correct? Appropriate? Authentic to my voice?” It helps to develop critical AI literacy as a 21st-century skill, just like grammatical competence (Godwin-Jones, 2022).

Leverage Mobile-First Design:

With almost all undergraduates in Iraq owning a smartphone, mobile-compatible AI tools (e.g., the Grammarly mobile app, Wordtune web app) should be incorporated into the writing curricula, and micro-writing tasks should be created to take advantage of access that is always available. This is consistent with MALL tenets and provides additional practice opportunities outside the classroom (Hwang et al., 2023).

Find Balance between Lower-Order and Higher-Order Skill Development:

Curricula for writing should be clear about the difference between editing (where AI excels) and composing (where human cognition is irreplaceable). Assignments could include:

- o Writing without AI to generate ideas

Editing with AI (for language editing)

- o What changed when I did XXX (to metacognition)

Create Institutional Guidelines to Use AI Responsibly:

Universities need to create extensive policies around AI-assisted writing that clearly define acceptable scaffolding (e.g. grammar checking) versus academic dishonest (e.g. topic-full text generation). Organising workshops for students and teachers also fosters responsible use in accordance with academic integrity standards (Zhou, 2025; Untsa & Nuha, 2025).

Support Instructor Professional Development:

AI-integrated pedagogy: Faculty training is necessary In its absence, even the best-intentioned tools may amplify inequity or entrench rote learning. A possible direction would be to pilot a “Digital Writing Pedagogy” module within teacher education programs at the University of Al-Qadisiyah.

Limitations of the Study



Although this research shows great promise, there are a few limitations that should be noted:

Sample Size and Generalizability:

Limitations of the study The study was conducted on 60 students only from one Iraqi university. While stratified sampling led to internal validity, results may not be applicable at different locations, proficiency levels, and educational systems.

Short-Term Intervention:

Intervention - 10-week intervention, Practical but may not measure longer-term effects of autonomy or creativity in writing A semester long or year long study would give further insight in to the retention and dependencies of these skills.

Tool-Specific Effects:

We centered on Grammarly and Wordtune in this study. The results of different AI tools (e.g. ChatGPT, QuillBot) may vary due to their algorithmic differences and feedback styles.

Self-Report Bias:

First, the Writing Self-Efficacy and Autonomy Scale is a self-reporting questionnaire, therefore subject to social desirability bias or inaccurate self-evaluation. Future studies could triangulate with behavioral measures (e.g., revision logs, eye-tracking).

Limited Use: Less Control Over Use in Home

Though there was oversight over in-class usage of AI, out-of-class usage with AI tools cannot be entirely controlled so as to rule out possible confounding variables.

Suggestions for Future Research

In order to expand this research and fill some gaps in the literature, we suggest the following research directions:

FF Long-Term Studies on AI Reliance

Follow students over a number of semesters to see if initial AI use causes the erosion of independent writing abilities or alternatively, fosters independent skill.



Tool Comparison Axis:

Aim : To examine the impact of different types of AI writing assistant (rule-based vs. generative, corrective vs. generative) on writing development in linguistic and rhetorical dimensions.

Instructor-Centered Research:

Investigate the perceptions, knowledge and practices of EFL educators on the incorporation of AI technology in the context of higher education in Iraq — an under explored yet crucial variable in studies focused on the students experience.

Step 13: Cross-Cultural Validation:

Implement this study in other Arabic-speaking EFL contexts (e.g., Jordan, Egypt, Sudan) to explore cultural and curricular moderators of AI effectiveness.

Interventions for AI Literacy

Develop and pilot clear instructional units that support students in critically considering when they should use, adapt, and/or incorporate the feedback from AI-produced writing feedback analyze AI not as a tool but rather as its own domain of literacy;

Neurocognitive Approaches:

Use objective methods such as keystroke logging or eye-tracking to identify cognitive load, revision behavior, and attention distribution when writing with AI.

In sum, the prospects of predictive writing AI technology to revolutionize EFL writing instruction at the University of Al-Qadisiyah can be a reality only if it is integrated with methodological insight, conscious conduct and cultural nuance. While Iraq builds its roadmap for 2030 to realize its educational vision for the country, the implementation of AI into foreign language learning should come not as a result of technological enthusiasm, but a commitment to develop independent, critical, and authentic writers.

6. References

Abuhussein, H., & Badah, A. (2025). The role of AI-based writing tools on L2 writing competency: Evidence from Palestinian EFL learners. *Dirasat: Human and Social Sciences*, 52(5), Article 6566. <https://doi.org/10.35516/hum.v52i5.6566>

Aladini, A., Ismail, S., Khasawneh, M., & Shakibaei, G. (2025). Self-directed writing development across computer/AI-based tasks: Unraveling the traces on L2



writing outcomes, growth mindfulness, and grammatical knowledge. *Computers in Human Behavior Reports*, 17, Article 100566. <https://doi.org/10.1016/j.chbr.2024.100566>

Alkamel, M., & Alwagieh, N. (2024). Utilizing an adaptable artificial intelligence writing tool (ChatGPT) to enhance academic writing skills among Yemeni university EFL students. *Social Sciences & Humanities Open*, 10, Article 101095. <https://doi.org/10.1016/j.ssaho.2024.101095>

Al-Raimi, M., Mudhsh, B., Al-Yafaei, Y., & Al-Maashani, S. (2024). Utilizing artificial intelligence tools for improving writing skills: Exploring Omani EFL learners' perspectives. *Forum for Linguistic Studies*, 6(2), 1–18. <https://doi.org/10.59400/fls.v6i2.1177>

Alshehri, M. (2025). Integrating automated writing evaluation into Saudi EFL students' writing practice. *Arab World English Journal*, 16(1), 68–82. <https://doi.org/10.24093/awej/vol16no1.5>

Altamimi, D. (2025). Unlocking potential: Saudi EFL male students' perspectives on AI tools for enhancing English writing proficiency. *Arab World English Journal (AI Special Issue)*, 34–51. <https://doi.org/10.24093/awej/ai.3>

Asadi, M., Ebadi, S., & Mohammadi, L. (2025). The impact of integrating ChatGPT with teachers' feedback on EFL writing skills. *Thinking Skills and Creativity*, 55, Article 101766. <https://doi.org/10.1016/j.tsc.2025.101766>

Ayoub, N., Joudi, N., Saba, M., & Saba, A. (2025). Integrating artificial intelligence to enhance writing proficiency: An exploratory study of EFL students' and instructors' perspectives at a university level in Lebanon. *European Scientific Journal, ESJ*, 21(14), 37–54. <https://doi.org/10.19044/esj.2025.v21n14p37>

Bensalem, E., Harizi, R., & Boujlida, A. (2024). Exploring undergraduate students' usage and perceptions of AI writing tools. *Global Journal of Foreign Language Teaching*, 14(2), 112–129. <https://doi.org/10.18844/gjflt.v14i2.9344>

Benson, P. (2011). *Teaching and researching autonomy* (2nd ed.). Pearson Education.

Budiyono, H. (2025). Exploring long-term impact of AI writing tools on independent writing skills: A case study of Indonesian language education students. *International Journal of Information and Education Technology*, 15(5), 230–237. <https://doi.org/10.18178/ijiet.2025.15.5.2306>



- Chen, B., Bao, L., Zhang, R., Zhang, J., Liu, F., Wang, S., & Li, M. (2023). A multi-strategy computer-assisted EFL writing learning system with deep learning incorporated and its effects on learning: A writing feedback perspective. *Journal of Educational Computing Research*, 61(1), 60–102. <https://doi.org/10.1177/07356331231189294>
- Dizon, G., & Gayed, J. (2021). Examining the impact of Grammarly on the quality of mobile L2 writing. *The JALT CALL Journal*, 17(2), 336–352. <https://doi.org/10.29140/jaltcall.v17n2.336>
- Fauziah, Z., & Minarti, S. (2023). Artificial intelligence-assisted writing application for improving writing skills of Islamic education students. *Jurnal Kependidikan Islam*, 13(2), 119–123. <https://doi.org/10.15642/jkpi.2023.13.2.119-123>
- Filfilan, S., & Alqurashi, H. (2025). The impact of using the QuillBot AI application on English writing skills development among EFL learners in Saudi Arabia. *Journal of English Language Teaching and Applied Linguistics*, 7(2), 234–248. <https://doi.org/10.32996/jeltal.2025.7.2.22>
- Godwin-Jones, R. (2022). Partnering with AI: Intelligent writing assistance and instructed language learning. *Language Learning & Technology*, 26(3), 10–26. <https://doi.org/10.64152/10125/73474>
- Guo, K., & Li, D. (2024). Understanding EFL students' use of self-made AI chatbots as personalized writing assistance tools: A mixed methods study. *System*, 125, Article 103362. <https://doi.org/10.1016/j.system.2024.103362>
- Guo, K., Pan, M., Li, Y., & Lai, C. (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, Article 100962. <https://doi.org/10.1016/j.iheduc.2024.100962>
- He, Y. (2024). The metaphor of AI in writing in English: A reflection on EFL learners' motivation to write, enjoyment of writing, academic buoyancy, and academic success in writing. *The International Review of Research in Open and Distributed Learning*, 25(3), 1–20. <https://doi.org/10.19173/irrodl.v25i3.7769>
- Hwang, W.-Y., Nurtantyana, R., Purba, S., Hariyanti, U., Indrihapsari, Y., & Surjono, H. (2023). AI and recognition technologies to facilitate English as foreign language writing for supporting personalization and contextualization in authentic contexts. *Journal of Educational Computing Research*, 61(4), 1008–1035. <https://doi.org/10.1177/07356331221137253>



- Khampusaen, D. (2025). The impact of ChatGPT on academic writing skills and knowledge: An investigation of its use in argumentative essays. *LEARN Journal: Language Education and Acquisition Research Network*, 18(1), 145–172. <https://doi.org/10.70730/pgcq9242>
- Lai, Z. (2025). The impact of AI-assisted blended learning on writing efficacy and resilience. *International Journal of Computer-Assisted Language Learning and Teaching*, 15(1), Article 377174. <https://doi.org/10.4018/ijcallt.377174>
- Liu, C., Hou, J., Tu, Y., Wang, Y., & Hwang, G.-J. (2021). Incorporating a reflective thinking promoting mechanism into artificial intelligence-supported English writing environments. *Interactive Learning Environments*, 31(9), 5614–5632. <https://doi.org/10.1080/10494820.2021.2012812>
- Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study. *Smart Learning Environments*, 11(1), Article 295. <https://doi.org/10.1186/s40561-024-00295-9>
- Mahmud, F. A. (2023). Investigating EFL students' writing skills through artificial intelligence: Wordtune application as a tool. *Journal of Language Teaching and Research*, 14(5), 1028–1037. <https://doi.org/10.17507/jltr.1405.28>
- Mansoor, H., Sumardjoko, B., & Sutopo, A. (2025). Attitudes of Pakistani undergraduate ESL students toward artificial intelligence in improving English writing skills. *World Journal of English Language*, 15(6), 369–384. <https://doi.org/10.5430/wjel.v15n6p369>
- Mohammed, S., & Khalid, M. (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(1), Article 343. <https://doi.org/10.1186/s40468-025-00343-2>
- Ngo, T., Chen, H., & Lai, K. (2022). The effectiveness of automated writing evaluation in EFL/ESL writing: A three-level meta-analysis. *Interactive Learning Environments*, 32(3), 727–744. <https://doi.org/10.1080/10494820.2022.2096642>
- Rad, H., Alipour, R., & Jafarpour, A. (2023). Using artificial intelligence to foster students' writing feedback literacy, engagement, and outcome: A case of Wordtune application. *Interactive Learning Environments*, 32(12), 5020–5040. <https://doi.org/10.1080/10494820.2023.2208170>
- Rahmi, R., Amalina, Z., Andriansyah, A., & Rodgers, A. (2024). Does it really help? Exploring the impact of AI-generated writing assistant on the students'



English writing. *Studies in English Language and Education*, 11(2), 587–605. <https://doi.org/10.24815/siele.v11i2.35875>

Rizkiani, S., Maulana, A., Resmini, S., Satriani, I., & Siliwangi, I. (2025). Examining the use of AI tools in academic writing: Effects on the critical thinking skills of EFL learners. *JELA (Journal of English Language Teaching, Literature and Applied Linguistics)*, 6(2), 155–170. <https://doi.org/10.37742/jela.v6i2.155>

Selim, A. (2024). The transformative impact of AI-powered tools on academic writing: Perspectives of EFL university students. *International Journal of English Linguistics*, 14(1), 14–28. <https://doi.org/10.5539/ijel.v14n1p14>

Syafei, M., & Nuraeningsih, N. (2025). Leveraging artificial intelligence in writing: ELT students' perspectives and experiences. *International Journal of Learning, Teaching and Educational Research*, 24(5), 389–408. <https://doi.org/10.26803/ijlter.24.5.22>

Takary, G., & Sabrina, S. (2025). Enhancing EFL writing proficiency through an AI-integrated platform: QuillBot's role in learners' engagement and skill development. In *Proceedings of the International Conference on Education* (Vol. 3, No. 1, pp. 3490–3502). <https://doi.org/10.32672/pice.v3i1.3490>

Untsa, F., & Nuha, U. (2025). Enhancing or inhibiting? The dual role of AI tools in EFL writing practices. *Esteem Journal of English Education Study Programme*, 8(2), 185–201. <https://doi.org/10.31851/esteem.v8i2.18553>

Van, N. (2025). The impact of AI tools on English writing skills development. *International Journal of Social Science Exceptional Research*, 4(3), 1–6. <https://doi.org/10.54660/ijsser.2025.4.3.01-06>

Wang, Z. (2020). Computer-assisted EFL writing and evaluations based on artificial intelligence: A case from a college reading and writing course. *Library Hi Tech*, 40(1), 80–97. <https://doi.org/10.1108/LHT-05-2020-0113>

Wei, P., Wang, X., & Dong, H. (2023). The impact of automated writing evaluation on second language writing skills of Chinese EFL learners: A randomized controlled trial. *Frontiers in Psychology*, 14, Article 1249991. <https://doi.org/10.3389/fpsyg.2023.1249991>

Widiati, U., Rusdin, D., Indrawati, I., Govender, N., & Mulyani, M. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. *Cogent Education*, 10(1), Article 2236469. <https://doi.org/10.1080/2331186X.2023.2236469>

Yannakoudakis, H., Andersen, Ø., Geranpayeh, A., Briscoe, T., & Nicholls, D. (2018). Developing an automated writing placement system for ESL learners. *Applied Measurement in Education*, 31(3), 251–267. <https://doi.org/10.1080/08957347.2018.1464447>

Zhao, D. (2024). The impact of AI-enhanced natural language processing tools on writing proficiency: An analysis of language precision, content summarization, and creative writing facilitation. *Education and Information Technologies*, 30, 8055–8086. <https://doi.org/10.1007/s10639-024-13145-5>

Zhao, X. (2022). Leveraging artificial intelligence (AI) technology for English writing: Introducing Wordtune as a digital writing assistant for EFL writers. *RELC Journal*, 54(3), 890–894. <https://doi.org/10.1177/00336882221094089>

Zhou, Y. (2025). Review of the studies on the impact of AI tools on EFL students writing. *Lecture Notes in Education Psychology and Public Media*, 2025, Article BO23872. <https://doi.org/10.54254/2753-7048/2025.bo23872>

Appendix A: Analytic Writing Assessment Rubric

(Printable Version – Score per dimension: 1–4; Total max = 16)

Dimension	4 (Excellent)	3 (Good)	2 (Fair)	1 (Poor)
1. Grammatical Accuracy & Mechanics	Few or no errors; structures consistently accurate and academic.	Occasional minor errors; mostly correct tense/agreement.	Frequent errors that occasionally hinder clarity.	Serious, pervasive errors; meaning often unclear.
2. Lexical Resource	Wide, precise vocabulary; effective collocations; minimal repetition.	Adequate vocabulary; some repetition or imprecision.	Limited range; basic/incorrect words; L1 interference.	Very restricted vocabulary; frequent misuse.
3. Coherence & Cohesion	Clear idea flow; strong paragraphs; effective linking devices.	Generally logical; some weak transitions; decent paragraphing.	Ideas lack connection; weak paragraphing; poor cohesion.	Disorganized; no structure; ideas disconnected.



Dimension	4 (Excellent)	3 (Good)	2 (Fair)	1 (Poor)
4. Task Achievement & Originality	Insightful, original ideas; clear voice; critical engagement.	Relevant ideas; some personal input; conventional.	Superficial or repetitive; minimal independent thought.	Off-topic/generic; no personal perspective; appears AI-generated.

AI Reliance Indicators (Posttest/Delayed Posttest Only)

Unnatural phrasing Register inconsistency Overuse of passive/nominalizations

Lack of personal stance Suspicious fluency (error-free but shallow)

Appendix B: Descriptive Statistics for Writing Performance

(Printable Table – Means and Standard Deviations)

Group	Time	Grammatical Accuracy	Lexical Resource	Coherence & Cohesion	Task Achievement & Originality	Total Score
Experimental	Pretest	2.10 (0.45)	2.03 (0.48)	2.17 (0.51)	2.07 (0.49)	8.37 (1.62)
	Posttest	3.23 (0.38)	3.10 (0.42)	3.07 (0.45)	2.83 (0.52)	12.23 (1.41)
	Delayed	3.00 (0.41)	2.93 (0.46)	2.90 (0.48)	2.60 (0.55)	11.43 (1.53)
Control	Pretest	2.13 (0.47)	2.07 (0.50)	2.20 (0.49)	2.10 (0.51)	8.50 (1.68)
	Posttest	2.47 (0.49)	2.40 (0.52)	2.53 (0.50)	2.33 (0.54)	9.73 (1.72)
	Delayed	2.40 (0.51)	2.33 (0.55)	2.47 (0.52)	2.27 (0.56)	9.47 (1.78)



Group	Time	Grammatical Accuracy	Lexical Resource	Coherence & Cohesion	Task Achievement & Originality	Total Score
)

Note: Scale per dimension = 1–4. Total max = 16.

Appendix C: Inferential Statistics – Group Comparisons

(Printable Table – Independent Samples t-Tests)

Time Point	Dimension	<i>t</i>	<i>df</i>	<i>p</i>	Cohen's <i>d</i>	Effect Size
Posttest	Total Score	6.82	58	<.001	1.76	Large
	Grammatical Accuracy	6.15	58	<.001	1.59	Large
	Lexical Resource	5.78	58	<.001	1.49	Large
	Task Achievement & Originality	3.21	58	.002	0.83	Large
Delayed Posttest	Total Score	5.24	58	<.001	1.35	Large
	Task Achievement & Originality	2.18	58	.033	0.56	Medium

Interpretation:

Small effect: $d \geq 0.2$ | Medium: $d \geq 0.5$ | Large: $d \geq 0.8$