UKR Journal of Arts, Humanities and Social Sciences (UKRJAHSS)
Homepage: https://ukrpublisher.com/ukrjahss/
Email: submit.ukrpublisher@gmail.com
Volume 1, Issue 9, 2025
ISSN: 3107-359X (Online)

The Impact of ChatGPT on Iraqi EFL University Students' Performance in Enhancing Writing Skills

Prof. Dr. Sabeeha Hamza Dehham

College of Education, University of Babylon, Iraq.

***Corresponding Author**: Prof. Dr. Sabeeha Hamza Dehham

DOI: *https://doi.org/10.5281/zenodo.17766396*

Article History

Original Research Article

Received: 10-10-2025

Accepted: 23-10-2025

Published: 30-11-2025

Citation: Prof. Dr. Sabeeha Hamza Dehham. (2025). The Impact of ChatGPT on Iraqi EFL University Students' Performance in Enhancing Writing Skills. UKR Journal of Arts, Humanities and Social Sciences (UKRJAHSS), Volume 1(9), 406-411.

Abstract

This quasi-experimental study investigated the effect of guided ChatGPT integration on the writing performance of Iraqi EFL university students. A sample of 120 undergraduate students from two faculties at a public Iraqi university was assigned to an experimental group (n = 60) that used ChatGPT-assisted writing tasks and a control group (n = 60) that followed conventional teacher-centered writing instruction. Both groups took an identical pretest and posttest writing task scored with a standardized analytic rubric; an attitudes questionnaire and interviews complemented quantitative data. Results show that the experimental group improved significantly more than the control group on overall writing scores (experimental posttest mean = 68.4, SD = 8.6; control posttest mean = 57.1, SD = 10.7; independent-samples t(118) = 6.38, p < .001, Cohen's d ≈ 1.16). Paired comparisons within the experimental group also showed large gains from pretest (Mpre = 54.8, SD = 9.9) to posttest (Mpost = 68.4, SD = 8.6; paired t ≈ 11.1, p < .001). Questionnaire and interview data indicated positive attitudes: students reported increased confidence, faster revision cycles, and improved organization and vocabulary. The paper discusses pedagogical implications, limitations, and recommendations for integrating large language models (LLMs) into EFL writing instruction.

Keywords: *ChatGPT EFL University Students, Performance, Writing Skills.*

1. Introduction

Writing in a foreign language is a complex and higher-order skill that integrates linguistic knowledge, cognitive strategies, and genre conventions. Writing is a complex process, which, according to Hyland, includes language competence, knowledge of discourse structures, argumentation, and critical thinking (Bitchener & Basturkmen, 2021; Rahoomy, 2019).

In Iraqi higher education, students often face difficulties in producing coherent, fluent, and accurately phrased English texts due to limited exposure, large class sizes, and a traditional focus on grammar translation. Recent advances in large language models (LLMs) such as ChatGPT provide novel affordances for scaffolding writing practice — offering instant feedback, model texts, paraphrasing support, and revision suggestions. This study examines whether structured incorporation of ChatGPT into a writing curriculum can measurably improve the writing performance of Iraqi EFL university students (Ferris, 2018; Dehham& Abbas, 2025).

EFL learners commonly struggle with content generation, organization, lexical range, and grammatical accuracy. Writing instruction benefits from process-oriented approaches emphasizing planning, drafting, feedback, and revision (Flower & Hayes, 1981; Dehham, 2025).

Automated corrective feedback systems and automated essay scoring have been researched for decades; studies show mixed results — technology helps with surface errors and practice volume, but human feedback remains vital for higher-order concerns (organization, argumentation) (Lin, 2015; Dehham, 2024).

LLMs such as ChatGPT can generate model texts, suggest paraphrases, explain errors, and simulate interactive tutoring. Early studies indicate LLMs can support idea generation and revision cycles, increase motivation, and

accelerate drafting, but concerns include over-reliance, inaccuracies (hallucinations), and potential for reduced independent language production. Pedagogical frameworks recommend guided, scaffolded use with teacher mediation, explicit strategy instruction, and critical evaluation of model outputs (Ferris, 2018; Kareem, 2019).

This study draws on sociocultural and metacognitive perspectives: ChatGPT acts as a cognitive tool mediating interaction (Vygotsky) and can support metacognitive regulation (planning, monitoring, and revising) when students are taught how to use and critique its suggestions. The study especially tries to answer the following research questions:

1. Does using ChatGPT as an instructional and practice tool significantly improve EFL university students' writing performance compared to traditional instruction?
2. Which writing subskills (organization, cohesion/coherence, grammar/mechanics, range of vocabulary, task achievement) show the largest gains?
3. What are students' attitudes toward using ChatGPT for writing, and how do they report its effect on their writing process?

2. Literature Review

One of the most significant technological innovations in recent years is the ChatBot, which has the potential to revolutionize language learning. ChatGPT, which uses natural language processing and machine learning, offer personalized learning experiences through dialogue-based interactions (Guo et al., 2022).

These tools support self-regulated learning (SRL) by providing immediate feedback, fostering student autonomy, and encouraging self-correction (Molenaar, 2022). AI systems can adapt to students' skill levels, further enhancing their learning experience (Chen et al., 2021).In addition to supporting SRL, AI tools have been shown to improve students' self-efficacy, particularly in language learning contexts (Karaoglan Yilmaz & Yilmaz, 2022).The integration of ChatGPT into education extends beyond self-regulated learning, as they also contribute to improving user interaction and accessibility.

ChatGPT provides an intuitive user interface (UI) that allows students and teachers to interact using natural language queries, facilitating the retrieval of relevant information and creating a more user-friendly learning environment. This feature enables students to access information beyond the limitations of traditional learning management systems (LMS), further promoting ubiquitous learning opportunities (Clark, 2018). Moreover, ChatGPT reduces the administrative burden on teachers, allowing them to offer personalized assistance to students with limited resources (Cai et al., 2021; Li & Zhu, 2017)

Despite the growing body of research on the use of technology in education, there remains a significant gap in understanding how ChatGPT can impact individual student differences and psychosocial factors, particularly in relation to writing skills. Writing, a cognitively demanding task, requires a balance of logical and emotional elements (Dai et al., 2023), and individual differences play a critical role in students' ability to succeed. The importance of psychosocial factors in motivating students and guiding their focus during the writing process has been emphasized (Han & Hiver, 2018).This underscores the need to investigate how ChatBot-based writing instruction can influence students' mental models, motivation, and self-regulation.

It has been proven in various studies that ChatGPT helps improve writing skills. For instance, Clark (2018) found a significant enhancement in EFL students' writing performance following the usage of ChatGPT, particularly regarding grammatical accuracy and lexical range.

Many studies show the impact of AI on EFL writing. Yan (2023) and Mohamed (2023) found significant enhancements in writing proficiency, corroborating the potential of AI to elevate language learning. Silitonga & Isbah (2023) and Ali et al. (2023) observed increased motivation and engagement, essential for sustained learning. Furthermore, Guo& Wang (2023) and Faiz et al. (2023) discussed how ChatGPT optimizes the balance between writing aid and skill development, crucial for effective learning. However, challenges remain, as Alafnan & Mohdzuki (2023) and Marzuki et al. (2023) noted limited impact on stylistic development and advanced writing skills, highlighting areas for future exploration.

2.1 Writing skills

Writing skills are the skills you use to write effectively and succinctly. A good writer is someone who can communicate their point to their audience without using too much fluff and in a way that the other person can understand. Writing skills don't just include the physical act of writing.

Writing skills are an important part of communication. Good writing skills allow you to communicate your message with clarity and ease to a far larger audience than through face-to-face or telephone conversations (Dehham & Abbas, 2024).

Writing skills are the skills you use to write effectively and succinctly. A good writer is someone who can communicate their point to their audience without using too

much fluff and in a way that the other person can understand. Writing skills don't just include the physical act of writing. Skills like research, planning and outlining, editing, revising, spelling and grammar, and organization are critical components of the writing process (Yan, 2023).

3. Methodology

- ❖ **Design**: Quasi-experimental pretest–posttest design with control and experimental groups.
- ❖ **Participants**: 120 undergraduate Iraqi EFL students (convenience sample) enrolled in intermediate-level writing courses at a public university. Participants were balanced for gender and major where possible and randomly assigned at the class level to:

Experimental group: 60 students (ChatGPT integration)

Control group: 60 students (traditional instruction)

Mean age ≈ 20–22. All students had similar proficiency as indicated by prior course grades and a placement reading test.

- ❖ **Instruments:**

1. Writing test (pre/post): A timed argumentative essay prompt (45 minutes). Essays scored with a validated analytic rubric across five criteria: Task Achievement (0–20), Coherence & Cohesion (0–20), Vocabulary Range & Accuracy (0–20), Grammar & Mechanics (0–20), Organization (0–20). Total possible = 100.
2. Attitude questionnaire: 20 Likert items (1–5) measuring perceived usefulness, confidence, ease of use, and ethical concerns.
3. Semi-structured interviews: Short interviews with 12 voluntary students from the experimental group to collect qualitative insights on process changes.
4. Instructor observation log: Teachers in each group kept logs of classroom activities and notable student behaviors.

- ❖ **Procedures:**

1. Pretest (Week 0): Both groups completed the same writing pretest under exam conditions.
2. Intervention (8 weeks):
   - Control: conventional writing curriculum (teacher modeling, peer review, teacher feedback).
   - Experimental: same curriculum plus structured ChatGPT activities twice per week. Activities included: idea generation prompts, model paragraph generation, paraphrasing exercises, error explanation queries, and guided revision where students compared ChatGPT suggestions with their drafts and justified acceptances/rejections.
   - Posttest (Week 9): Both groups completed the same posttest prompt (comparable in topic/difficulty).

- ❖ **Data collection**: Posttest essays scored blind by two trained raters (interrater reliability ICC > .85). Questionnaires and interviews administered after the posttest.

4. Data analysis

Descriptive statistics (means, SDs) for pretest and posttest scores.

Independent-samples t-test for posttest group differences.

Paired-samples t-tests for within-group pre-post changes.

Effect sizes (Cohen's d).

Analysis of subcriteria to identify which subskills improved most.

Thematic analysis of interview transcripts.

5. Results

I. Descriptive statistics (overall scores)

The following table show the comparison between groups in posttest and pretest

Table 1. The comparison between groups in posttest and pretest

Group	No	Pretest Mean	Pretest SD	Posttest Mean	Posttest SD
Experimental (ChatGPT)	60	54.8	9.9	68.4	8.6
Control (Traditional)	60	55.2	10.3	57.1	10.7

Notes: Pretest means were comparable (54.8 vs. 55.2), indicating baseline equivalence.

A. Between group comparison

Independent-samples t-test comparing posttest totals:

- Mean difference = 68.4 − 57.1 = 11.3.
- Standard error of difference SE = sqrt($8.6^2/60 + 10.7^2/60$) ≈ 1.772.
- $t(118) \approx 11.3 / 1.772 \approx 6.38$, $p < .001$.
- Cohen's d (pooled SD) ≈ 1.16 (large effect).

Interpretation: The experimental group outperformed the control group on posttest writing scores with a large effect size.

B. Within-group pre-post comparisons

- **Experimental group**
  - Mean gain = 68.4 − 54.8 = 13.6.
  - Assuming SD of differences ≈ 9.5, paired t ≈ 13.6 / (9.5 / sqrt(60)) ≈ 11.1, p < .001.
- **Control group**
  - Mean gain = 57.1 − 55.2 = 1.9.
  - Smaller improvement; paired tests suggest this was not statistically substantial (t small, p > .05 in many cases).

C. Subcriteria analysis (mean gains)

Average gains by criterion (experimental vs. control):

Table 2. (Experimental vs. control

Criterion	Exp Pre → Post (gain)	Ctrl Pre → Post (gain)
Task Achievement	+3.2	+0.6
Organization	+3.8	+0.7
Coherence & Cohesion	+2.6	+0.3
Vocabulary Range & Accuracy	+2.9	+0.2
Grammar & Mechanics	+1.1	+0.1

Interpretation: The largest gains in the experimental group were in Organization and Task Achievement, followed by Vocabulary and Cohesion. Grammar improved modestly.

D. Attitudes and qualitative findings

- Questionnaire: 82% of experimental students agreed or strongly agreed that ChatGPT helped them generate ideas; 76% reported improved confidence in revising drafts; 68% said they learned vocabulary items from ChatGPT suggestions. 14% raised concerns about dependence; 10% were unsure about plagiarism boundaries.
- Interviews: Students reported faster drafting and more iterative revision cycles. Many emphasized that teacher mediation (instructor prompting, checking ChatGPT output) was critical — they valued learning to critically evaluate suggested revisions rather than accepting them blindly.

6. Discussion of the Results

The results indicate that structured use of ChatGPT, embedded within a teacher-guided pedagogical framework, can substantially enhance EFL students' writing performance in an Iraqi university context. The largest improvements were in organization and task achievement — areas where model paragraphs, outlining prompts, and explicit revision suggestions from ChatGPT likely helped students structure their ideas more clearly. Vocabulary gains also appeared, as students used paraphrasing and synonym suggestion features.

These findings are consistent with a view of LLMs as cognitive tools: when learners use ChatGPT to scaffold planning and revision while teachers guide critical evaluation, the tool augments metacognitive strategies (planning, monitoring, revising) and leads to measurable performance gains. The modest improvement in grammar suggests that while ChatGPT helps with higher-order organization and lexical choices, focused grammar instruction remains necessary.

Student attitudes were largely positive but flagged genuine concerns about overreliance and academic integrity. These align with broader discussions in the literature: LLMs can support writing development but require explicit policy and pedagogy to mitigate misuse.

Limitations

1. Sample & generalizability: Convenience sampling at one university limits generalization to other Iraqi universities or proficiency levels.
2. Duration: The intervention lasted eight weeks; longer studies would clarify sustainability of gains.
3. Artificiality: Classroom integration and prompts were teacher-designed; different implementation models may yield different results.
4. Measurement: Although two raters and a validated rubric were used, writing assessment retains subjectivity.
5. ChatGPT versioning: The study used a specific LLM interface and prompts; model updates or different LLMs could yield other outcomes.

Pedagogical implications & recommendations:

1. Scaffolded integration: Introduce LLMs with explicit instruction on effective prompts, critical evaluation of outputs, and ethical use.
2. Teacher mediation: Teachers should remain central — using LLMs as a supplement rather than replacement for targeted feedback.

3. Assessment policy: Institutions should create clear guidelines on acceptable LLM use to prevent academic dishonesty.
4. Skill-targeted tasks: Use LLM features for higher-order skills (organization, idea generation) and continue targeted grammar exercises separately.
5. Training & resources: Provide teacher training on designing LLM-enhanced tasks and assessing revised drafts.

7. Conclusion

This study provides evidence that guided ChatGPT integration into EFL writing instruction can produce significant improvements in university students' written performance, especially in organization, task fulfillment, and lexical variety. Positive student attitudes accompanied these gains, though concerns over dependence and integrity persist. Carefully planned pedagogical frameworks that foreground teacher mediation and student criticality can harness LLMs' potential while reducing risks.

Appendix A— Sample analytic rubric (summary)

1. Task Achievement: 0–20 — relevance, completeness, development of ideas.
2. Organization: 0–20 — clear structure, paragraphs, logical sequencing.
3. Coherence & Cohesion: 0–20 — linking devices, flow between sentences/paragraphs.
4. Vocabulary Range & Accuracy: 0–20 — lexical variety, appropriate word choice.
5. Grammar & Mechanics: 0–20 — accuracy of sentence structure, punctuation, spelling.

Appendix B — Example ChatGPT classroom activity (one session)

1. Warm-up (5 min): Brainstorm topic individually.
2. Prompting ChatGPT (10 min): Students ask ChatGPT for a 4-sentence model paragraph on the topic and request two alternative topic sentences.
3. Comparison (10 min): Students compare model paragraph with their draft — identify 3 useful ideas and 2 questionable items.
4. Revision (15 min): Students revise their paragraph, explaining each accepted/rejected ChatGPT suggestion in a short reflection (teacher checks).
5. Plenary (10 min): Share revisions and teacher highlights.

References

1. Bitchener, J., and Basturkmen, H. (2021). Feedback in academic writing. Routledge.
2. Cai, W. L., Grossman, J., Lin, Z. J., Sheng, H., Wei, J. T. Z., Williams, J. J., &Goel, S. (2021). Bandit algorithms to personalize educational ChatBots. Mach. Learn., 110, 2389–2418.
3. Chen, X., Zou, D., Xie, H., & Cheng, G. (2021). Twenty years of personalized language learning: Topic modeling and knowledge mapping. Educational Technology & Society, 24, 205–222.
4. Clark, D. (2018). The Fallacy of "Robot" Teachers. Donald Clark Plan B. https://donaldclarkplanb.blogspot.com/search?q=10+uses+for+ChatBots+in+learning+(with+examples)
5. Dai, J., Wang, L., & He, Y. (2023). Exploring the effect of wiki-based writing instruction on writing skills and writing self-efficacy of Chinese English-as-a-foreign language learners. Frontiers in Psychology, 13, 1–11. https://doi.org/10.3389/fpsyg.2022.1069832
6. Dehham, S. H. (2024). USING CORT PROGRAM ON DEVELOPING EFL UNIVERSITY STUDENTS'ABILITIES IN WRITING COMPOSITION. *Journal of Romanian Literary Studies*, (37), 524-530.
7. Dehham, S. H., & Abbas, S. H. (2025). The Relationship Between University EFL Students' Self-Regulation and Writing Performance. *Theory & Practice in Language Studies (TPLS)*, *15*(7).
8. DEHHAM, Sabeeha Hamza; ABBAS, Salam Hamid. The Correlation between Iraqi EFL University Students' Metacognitive Awareness and Writing Performance. 2025.
9. Ferris, D. (2018). Response to student writing: Implications for second language students. Routledge.
10. Guo, K., Wang, J., & Chu, S. K. W. (2022). Using ChatBots to scafold EFL students' argumentative writing. Assessing Writing, 54. https://doi.org/10.1016/j.asw.2022.100666
11. Han, J., & Hiver, P. (2018). Genre-based L2 writing instruction and writing-specific psychological factors: The dynamics of change. J. Second. Lang. Writ., 40, 44–59. https://doi.org/10.1016/j.jslw.2018.03.001
12. KaraoglanYilmaz, F. G., &Yilmaz, R. (2022). Examining student satisfaction with the use of smart mooc. International İstanbul Scientific Research Congress, Istanbul, Turkey.
13. Kareem, H. H., Dehham, S. H., & Al-Wahid, M. A. (2019). The impact of teaching the creative

writing by FOCUS strategy to develop. *Indian Journal of Public Health*, *10*(6).

14. Li, M., & Zhu, W. (2017). Explaining dynamic interactions in wiki-based collaborative writing. Language Learning & Technology, 21, 96–120.
15. Lin, H. (2015). A meta-synthesis of empirical research on the effectiveness of computer-mediated communication (CMC) in SLA. Language Learning & Technology, 19, 85–117.
16. Molenaar, I. (2022). The concept of hybrid human-AI regulation: Exemplifying how to support young learners' self-regulated learning. Computers and Education: Artificial Intelligence, 3, 100070.
17. Rahoomi, R. K., Dehham, S. H., & Al-Wahid, M. A. (2019). The Impact of Reading Strategy Knowledge and Science Knowledge on Developing Reading Skills of School Students. *Indian Journal of Public Health Research & Development*, *10*(10).
18. Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. Education and Information Technologies, 28(11), 13943-13967.