



Effects of DeepSeek-assisted Writing Instruction on EFL Learners' Writing Performance

Shuai Hu*

College of Foreign Languages
Bohai University
Jinzhou, Liaoning, China
hushuai6@163.com

Lei Wang

College of Foreign Languages
Bohai University
Jinzhou, Liaoning, China
2629716424@qq.com

Abstract

To investigate the effects of DeepSeek assisted writing instruction on different dimensions of learners' English writing skills, a prospective, quasi-experimental study was conducted in 2025. Employing a mixed research design, the study involved 31 students over six weeks. The experiment consisted of three stages: baseline period, intervention period and follow-up period. During the intervention period DeepSeek was employed to assist EFL learners' writing. Learners' writing production of each period was collected and rated in a randomized fashion to minimize grading bias. Then a semi-structured interview was conducted to further elucidate proficiency-dependent perceptions. Results revealed that DeepSeek significantly enhanced micro-level writing skills (i.e., vocabulary, grammar, mechanics) across all proficiency groups. However, higher-proficiency learners demonstrated greater gains in macro-level performance, while lower-proficiency learners exhibited over-reliance on technical corrections. Thematic analysis highlighted divergent attitudes: advanced learners valued DeepSeek's in-depth logical reasoning scaffolds, whereas lower-proficiency learners valued error reduction and were confounded by cognitive over-load. These findings underscore the necessity of proficiency-tailored AI assistance together with teacher guidance and peer collaboration in EFL writing pedagogy.

CCS Concepts

- Applied computing; • Education; • Computer-assisted instruction;

Keywords

DeepSeek, Artificial intelligence, AI-assisted writing, EFL Writing Performance

ACM Reference Format:

Shuai Hu and Lei Wang. 2025. Effects of DeepSeek-assisted Writing Instruction on EFL Learners' Writing Performance. In *The 2nd Guangdong-Hong Kong-Macao Greater Bay Area Education Digitalization and Computer Science International Conference (EDCS 2025), April 18–20, 2025, Shenzhen, China*. ACM, New York, NY, USA, 6 pages. <https://doi.org/10.1145/3746469.3746577>

*Corresponding author.



This work is licensed under a Creative Commons Attribution 4.0 International License.
EDCS 2025, Shenzhen, China

© 2025 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-1381-1/2025/04
<https://doi.org/10.1145/3746469.3746577>

1 Introduction

In the field of second language acquisition, writing is one of the most complex and challenging skills [1]. The development of proficient writing skills is integral to the academic and professional success of university students [2]. Currently, there has been a growing trend in the use of chatbots within the realm of English as a Foreign Language (EFL) instruction [3]. And an increasing number of research has affirmed the transformative potential of GenAI chatbots on learner writing performance, suggesting that regular and iterative feedback can substantially improve writing outcomes [4]. The proliferation of DeepSeek R1 by the end of 2023 has once again propelled the discourse on the application of AI in foreign language education which was sparked by the launch of ChatGPT to unprecedented heights. Many researchers have confirmed the effectiveness of GenAI chatbots in second language writing practice [5] [6], however, most of these studies are based on the application of ChatGPT series. As an AI agent that is powerful in fluency and practicality of natural language interaction, the impact of DeepSeek on EFL writing practice needs to be assessed empirically to discover the possibility of exploiting it to assist Chinese EFL learners in English writing.

2 Research questions

The present study aims at examining the effects of DeepSeek-assisted writing instruction on EFL learners' achievement in general writing performance and its different dimensions and how initial language proficiency mediates the effects. To meet the mentioned objectives, the following research questions were addressed in this study.

RQ1: What are the effects of DeepSeek-assisted writing instruction have on EFL learners' overall writing performance and different dimensions of writing?

RQ2: What are the factors that influence the ultimate writing performance of EFL learners?

3 Research Design

3.1 Design of the research

The current study employed a mixed-methods quasi-experimental design, incorporating both quantitative analysis (i.e. MANOVA test, t-test) and qualitative analysis (i.e. interview). The independent variables in this study are the different writing instruction methods, while the dependent variables are the writing scores and learner perceptions. As shown in Figure 1, a single-group repeated measures design was adopted for this type of research design is credited for its efficiency in control for inter-individual differences,

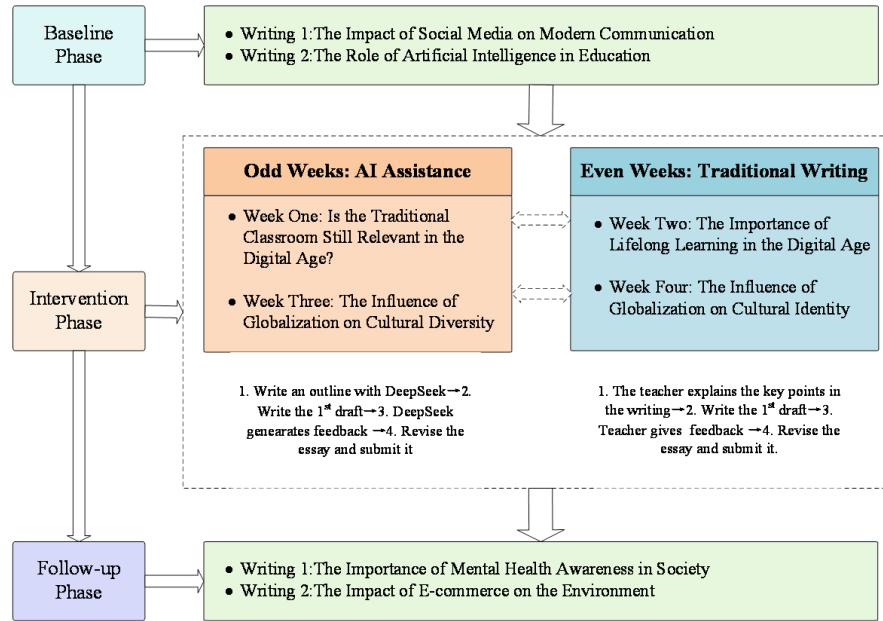


Figure 1: Flowchart of the research design.

thereby reducing error variance and enhancing the statistical power of hypothesis testing [7].

3.2 Participants

In this study, a total of 31 college students majoring in English at one university in north China, taking advanced English course in their 3rd year, were the participants in the study. They had four class periods of advanced English course each week. 83.33% of them had passed TEM-4. The participants were uninformed of the impending study and remained oblivious to DeepSeek's role in revision support. Consequently, they were unaware of the research context during their initial writing phase.

3.3 Procedures

The experiment was conducted over three distinct phases: baseline period, intervention period, and follow-up period. During the baseline phase (one weeks), participants were asked to submit two writing samples to establish individual writing proficiency baselines. They were asked to write the essays in class within 30 minutes all by themselves without resorting to any AI tool or other forms of writing assistants. Then came the intervention phase which spanned four weeks, with odd-numbered weeks designated for AI-assisted writing using DeepSeek and even-numbered weeks for traditional writing. Then it was the follow-up phase, in which two unassisted writing samples were collected to assess changes in learners' writing performance. The sets of writing tasks in odd-numbered weeks and even-numbered weeks are semantically related but do not overlap in content. To ensure rigorous control, a double-blind procedure was employed, with essays scored by evaluators unaware of the writing mode or timing.

A multi-dimensional comparison approach was adopted. Horizontal comparisons examined quality differences between

DeepSeek-assisted and traditional writing on the similar topics. Vertical comparisons traced individual development trajectories from baseline to intervention and follow-up phases. Statistical methods included repeated measures MANOVA to compare scores across phases. To ensure reliable research findings, a semi-structured interview was conducted by the end of the follow-up period.

3.4 Data Collection

3.4.1 Writing rating rubric. The rubric used to evaluate students' writing performance encompasses three five dimensions: content, organization, vocabulary, grammatical accuracy, and mechanics. This rubric draws inspiration from scales proposed by Li et al. [8]. The content dimension evaluates the relevance and appropriateness of the argument to the topic. The organization dimension measures the fluency and coherence of the components in writing and the clarity of the underlying logic. The vocabulary evaluates the range, sophistication and appropriateness of words. The accuracy dimension measures the grammatical accuracy, and the mechanics dimension evaluates the mastery of spelling, punctuation and capitalization. Each dimension is categorized into four levels, with a maximum score of 10. The overall writing performance is calculated as the weighted sum of the five dimensions: content (30%), organization (20%), vocabulary (20%), accuracy (20%) and mechanics (10%).

3.4.2 Writing rating. The evaluation of student's writing was conducted by two experience teachers in the same college. If there is a large difference in the scores for each dimension (i.e., a discrepancy of ± 3 points or more), a third rater will be brought in to mediate and help reach a consensus. A reliability test was conducted on the scores provided by the two initial raters using Cohen's kappa, and the Cohen's kappa value was 0.67, which indicates a substantial

Table 1: The Scores in Baseline Phase and Intervention Phase.

	Phase	Mean	Std. Deviation	N
Content	Baseline	5.67	.835	31
	Intervention	6.38	.769	31
Organization	Baseline	6.12	.681	31
	Intervention	6.61	.850	31
Vocabulary	Baseline	5.98	.781	31
	Intervention	7.89	.732	31
Accuracy	Baseline	5.82	.875	31
	Intervention	8.15	.682	31
Mechanics	Baseline	6.87	1.01	31
	Intervention	8.35	.772	31
Score	Baseline	30.46	3.870	31
	Intervention	37.37	3.312	31

level of agreement between the raters, suggesting that the scoring process was reliable.

3.4.3 Interview. Students were grouped according to their TEM-4 scores (Test for English Majors-Band 4). The TEM-4 scores (range: 38–85; $M = 63.23$, $Mdn = 63.5$, $SD = 10.65$) were found to be normally distributed based on the Shapiro-Wilk test ($W = 0.952$, $*p^* = 0.132$). Hence, the top 8 students (25%) were assigned to the high-proficiency group (scores 72–85), and the bottom 8 students (25%) to the low-proficiency group (scores 38–56). A semi-structured interview with 6 participants (3 from high proficiency group, 3 from low-proficiency group) was conducted by the end of the follow-up period. The interview consisted of the following questions: 1. Do you think DeepSeek assistance improved your writing performance or not? If yes, which aspect of writing did it influence you the most? 2. What advantages does the DeepSeek-assisted writing approach have compared to regular writing approach? 3. What are the challenges that you met with DeepSeek assisted writing practice?

3.4.4 Data Analysis. For the quantitative data, we employed SPSS 27 to carry out Independent T-test and MANOVA analysis. Independent T-test is employed to test whether significant differences exist between the writing scores of baseline phase and intervention phase, and those of the baseline phase and the follow-up phase. The aim was to assess the impact of DeepSeek mediation on students' writing performance, while considering the pre-existing writing levels. The qualitative data of the interview was organized thematically, aligned with the three interview questions, to highlight students' perceived benefits, challenges, and the specific dimensions influenced by the tool.

4 Results

4.1 Differences in the writing scores in three phases

As can be seen from Table 1, the total scores and those of each dimension after the application of DeepSeek as a writing assistant were improved at the end of the intervention phase. And whether the differences were significant should be further examined through

Independent T-test, as shown in Table 2. The P-values of content, vocabulary, accuracy, mechanics and the total score are all less than 0.05, which indicates significant differences between learner writing performance in the baseline phase and the intervention phase. While to test whether the positive effect was short-term effect which would fade within a short period of time, the comparison between learners' initial scores and the scores in the follow-up phase was also conducted, as shown in Table 2. It is obvious that the scores in the follow-up phase were higher than those in the baseline period.

According to Table 3., the P-values of dimensions such as content, vocabulary, accuracy and mechanics are all less than 0.05, which indicates significant differences in aforementioned dimensions in the baseline phase and the follow-up phase. However, for the dimension of organization, the P-value is 0.882, which largely exceeds 0.05, indicating that the variances are not equal in this very dimension, which means that no significant difference is detected in organization between the baseline phase and the follow-up phase.

4.2 Factors influencing the ultimate writing performance

In the present study, we first establish a model that only includes baseline scores and analyze how initial writing ability explains differences in writing performance, then, add intervention measures as a new variable based on the existing model, and analyze how much additional variance can be explained by the intervention measures. This hierarchical modeling can clarify the independent contribution of each variable.

The overall model was tested using multivariate analysis of variance (MANOVA) to investigate the factors that impact the five dimensions of writing performance. Before the test was administered, the correlation between the dependent variables was first tested and the results showed that each dimension had the same variance covariance, which ensures the validity and reliability of MANOVA test. Table 4 shows the results of multivariate tests.

To begin with, the results indicate a highly significant effect of baseline score on the combined dimensions (Multivariate Tests: Pillai's Trace = 0.875, $F = 77.038$, $df = 5$, $p < 0.001$, Partial $\eta^2 = 0.875$). Partial η^2 represents the proportion of variance explained

Table 2: Independent T-test of the Scores in Baseline Phase and Intervention Phase.

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Content	Equal variances assumed	1.219	.288	-2.786	14	.015	-.60000	.21536	-1.0619	-.13807
	Equal variances not assumed			-2.786	12.896	.016	-.60000	.21536	-1.0656	-.13434
Organization	Equal variances assumed	.105	.751	-3.879	14	.002	-.82500	.21265	-1.2811	-.36889
	Equal variances not assumed			-3.879	12.857	.002	-.82500	.21265	-1.2849	-.36506
Vocabulary	Equal variances assumed	1.198	.292	-17.527	14	.000	-2.0250	.1155	-2.2728	-.17772
	Equal variances not assumed			-17.527	12.976	.000	-2.0250	.1155	-2.2746	-.17754
Accuracy	Equal variances assumed	1.143	.303	-16.168	14	.000	-2.1375	.1322	-2.4210	-.18540
	Equal variances not assumed			-16.168	13.993	.000	-2.1375	.1322	-2.4211	-.18539
Mechanics	Equal variances assumed	.162	.694	-4.433	14	.001	-1.2125	.2735	-1.7991	-.6259
	Equal variances not assumed			-4.433	13.144	.001	-1.2125	.2735	-1.8027	-.6223
Total	Equal variances assumed	.857	.370	-9.873	14	.000	-6.8000	.68874	-8.2772	-.53227
	Equal variances not assumed			-9.873	11.023	.000	-6.8000	.68874	-8.3155	-.52844

Table 3: Independent T-test of the Scores of Baseline and Follow-up Phases.

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	Content	5.921a	1	5.921	8.493	.005	.124
	Organization	.017b	1	.017	.022	.882	.000
	Vocabulary	8.180c	1	8.180	13.649	.000	.185
	Accuracy	9.524d	1	9.524	11.789	.001	.164
	Mechanics	7.863e	1	7.863	8.301	.005	.122
Intercept	Content	2219.537	1	2219.537	3183.510	.000	.982
	Organization	2309.094	1	2309.094	2990.466	.000	.980
	Vocabulary	2495.171	1	2495.171	4163.427	.000	.986
	Accuracy	2394.453	1	2394.453	2963.866	.000	.980
	Mechanics	3235.427	1	3235.427	3415.556	.000	.983

uniquely by the independent variable. According to Cohen's (1988) guidelines, 0.875 indicates large effect. Hence, this means that initial writing proficiency accounts for 87.5% of the variance across the five dimensions. On the other hand, the effect of the AI intervention on the overall writing performance was also examined. The results

revealed a significant impact on students' writing performance, independent of baseline proficiency (Multivariate Tests: Pillai's Trace = 0.675, F = 22.872, df = 5, p < 0.001, Partial η^2 = 0.675). This indicates that the AI assistance also contributes significantly to the

Table 4: Impact of Baseline Writing Proficiency on Writing Improvement in Follow-up Phases.

Multivariate Tests ^a		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power ^c
Effect									
Intercept	Pillai's Trace	.147	1.902b	5	55	.109	.147	9.512	.601
	Wilks' Lambda	.853	1.902b	5	55	.109	.147	9.512	.601
	Hotelling's Trace	.173	1.902b	5	55	.109	.147	9.512	.601
Baseline Score	Pillai's Trace	.875	77.038b	5	55	.000	.875	385.188	1.000
	Wilks' Lambda	.125	77.038b	5	55	.000	.875	385.188	1.000
	Hotelling's Trace	7.003	77.038b	5	55	.000	.875	385.188	1.000
AI Intervention	Pillai's Trace	.675	22.872b	5	55	.000	.675	114.362	1.000
	Wilks' Lambda	.325	22.872b	5	55	.000	.675	114.362	1.000
	Hotelling's Trace	2.079	22.872b	5	55	.000	.675	114.362	1.000

improvement of students' writing skills, explaining 67.5% of the variance in writing performance.

4.3 Results of the interview

The semi-structured interviews with 6 participants revealed three overarching themes regarding AI chatbots' impact on EFL writing instruction: (i) generally perceived improvement in engagement and writing production, (ii) mixed attitudes towards the impact of DeepSeek on different dimensions of writing performance, and (iii) challenges and limitations of DeepSeek as a writing assistant.

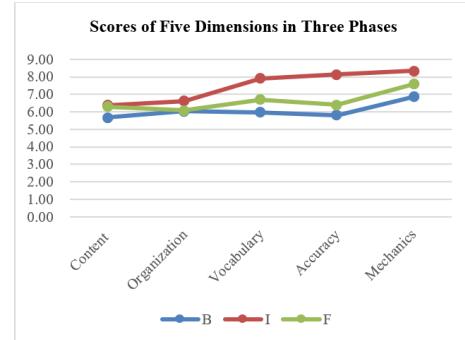
All the interviewees reported that DeepSeek enhanced their confidence and performance in EFL writing, with improvements concentrated in grammar, vocabulary, and structural coherence. 66.67% (n=4) of respondents highlighted lexical enhancement as the most influential feature, particularly for academic terms, Participant 6 remarked. Grammar correction was also frequently cited. One student noted, "DeepSeek caught subtle tense errors I'd repeated for years—finally, I understand where I go wrong" (student 3). Approximately 50% of advanced learners emphasized improvements in content. However, organization was rarely mentioned as an area of growth, especially students with lower-level of proficiency.

As for the advantages of DeepSeek over regular writing approach, 83.33% (n=5) participants valued the immediacy and personalization of feedback, contrasting it with delayed teacher evaluations. Advanced learners appreciated the DeepSeek's feedback comprehensiveness and meticulousness. Despite these benefits, students also reported some challenges, for instance, cognitive over-load, as lamented by student 5, "Sometimes it gave me too many suggestions, and I just couldn't quite digest them all, especially when it came to organization or structure." And student 6 also complained about the complexity of the feedback suggestions. Besides, 3 of the interviewed students expressed need for guidance and collaboration.

5 Discussion

5.1 General impacts on students' writing performance

The results show that the intervention of DeepSeek as an EFL writing assistant has significant effects upon students' overall writing

**Figure 2: Scores of Five Dimensions in Three Phases.**

performance. This echoed with Guo et al.'s research [9] and Tsai et al.'s study [10]. Figure 2 demonstrates the developmental curve of EFL learners' writing performance across three experimental phases.

As can be seen from Figure 2, during the Baseline phase (B), scores across all dimensions exhibit moderate initial with mechanics and accuracy showing slightly lower averages. A noticeable upward trend emerges in the Intervention phase (I), particularly in organization and vocabulary, which demonstrate steep improvements, likely attributable to the AI mediation. Accuracy and Mechanics exhibit more gradual but consistent gains. In the Follow-up phase (F), scores stabilize for most dimensions, suggesting retention of acquired skills. Organization maintains its elevated performance, which is in line with Bai & Wei's study, while vocabulary shows a minor decline, possibly indicating variability in long-term lexical retention. Accuracy and mechanics continue to improve marginally, reflecting sustained refinement in technical aspects of writing, which corresponds with Tam's study [11]. And this mirrors the overall positive pattern of development, peaking during the intervention and stabilizing post-intervention, which highlights the DeepSeek intervention's efficacy in enhancing writing competencies, with technical skills requiring prolonged reinforcement.

Results of the interview suggest that students had difficulty taking in the suggestions about organization of their compositions, and this result is consistent with Zhou et al. [12], Guo & Li [13] and

Li & Huang [14], which indicates challenges for EFL learners to digest and incorporate AI suggestions in macro aspect of writing into their writing practice. One of the key factors in improving text structure performance is to enhance students' accurate understanding of multi-level concept logic relationships related to the topic, and this is more demanding than adjust the problems in mechanics or lexical choices.

5.2 Factors influencing the ultimate writing improvement

The results of MANOVA analysis underscore two pivotal factors: initial writing proficiency and AI-driven instructional intervention. To start with, baseline scores exerted a highly significant effect. This highlights the critical role of students' pre-existing language ability in facilitating themselves in incorporating AI assistance with their writing practice. Besides, the DeepSeek intervention demonstrated a substantial independent impact on the overall writing performance [15–17]. The intervention consistently enhanced writing skills across proficiency levels. Sustained AI support facilitated progressive improvements in accuracy and mechanics, though interviews revealed proficiency-dependent perceptions of the tool's efficacy. Less advanced learners tended to confounded by the cognitive over-load caused by the excessive abundance and complexity of Deep Seek feedbacks which is in line with AlAfnan's research [18].

These effects suggest a dual mechanism: initial proficiency lays the foundation for learning potential, while the DeepSeek intervention serves as a catalyst to unlock incremental improvements, even for students with lower baseline skills. The thematic analysis of semi-structured interviews with students with different baseline writing proficiency also supported these findings.

6 Conclusion

In summary, this study focuses on the impact of DeepSeek assistance on different dimensions of EFL learners' writing performance, and this study verifies the effectiveness of AI-assisted writing approach, especially in terms of micro-level writing skills, although no significant influence was detected in macro-level skill such as organization. The Learners' initial proficiency and the DeepSeek intervention both mediated various dimensions of learners' final writing performance, and students with different proficiency levels demonstrated different development patterns. The findings of the present research highlight the need for teachers to exploit AI chatbots to provide differentiated instructions and tailored interventions based on the baseline proficiency and the necessity of instructions and interaction between teachers and students. Future research is expected to measure the long-term effects of AI intervention on learners' writing competence through a longitudinal study design and to explore interaction mechanism between baseline proficiency and intervention intensity, as well as the longitudinal sustainability of AI-driven gains across diverse educational contexts.

Acknowledgments

This work was supported by Bohai University, Research and Practice on the Practical and Innovative Competency Development System for Professional Master's Students in English Subject Pedagogy

(YJG20240013); Study on Deep Integration Between English Reading Teaching and Information Technology in Senior High Schools (2024ZXXJG05); Research on AI-Empowered Strategies for Rural Foreign Language Teacher Training. Special gratitude is extended to the anonymous reviewers for their constructive feedback.

References

- [1] Pooja Anggusari, Sri Wahyuni. 2023. Direct Written Corrective Feedback for Tenth Graders Recount Text: Adequate Practice to Boost Sentential Accuracy. *Journal of English Teaching*. 9, 310–322. <https://doi.org/10.33541/jet.913.4934>
- [2] Rana AlGhamdi. 2024. Exploring the impact of ChatGPT-generated feedback on technical writing skills of computing students: A blinded study. *Educ. Inf. Technol.* 29, 18901–18926. <https://doi.org/10.1007/s10639-024-12594-2>.
- [3] Min Hwang, Robert Jeens, and Hye-Kyung Lee. 2024. Exploring Learner Prompting Behavior and Its Effect on ChatGPT-Assisted English Writing Revision. *Asia-Pacific Edu. Res.* <https://doi.org/10.1007/s40299-024-00930-6>.
- [4] Victor Zhang and Ken Hyland. 2023. The role of digital literacy in student engagement with automated writing evaluation (AWE) feedback on second language writing. *Comput. Assist. Lang. Learn.* 1–26. <https://doi.org/10.1080/09588221.2023.2256814>.
- [5] Lifeng Bai and Yujie Wei. 2024. Exploring EFL Learners' Integration and Perceptions of ChatGPT's Text Revisions: A Three-Stage Writing Task Study. *IEEE Trans. Learn. Technol.* 17, 2161–2172. <https://doi.org/10.1109/TLT.2024.3491864>.
- [6] Chen Wang. 2024. Exploring Students' Generative AI-Assisted Writing Processes: Perceptions and Experiences from Native and Nonnative English Speakers. *Tech Know Learn.* <https://doi.org/10.1007/s10758-024-09744-3>.
- [7] Ibrahim Abba Mohammed, Oluwole Caleb Falode, Ibrahim Ismaila Kuta & Ahmed Bello. 2024. Effect of game-based learning on educational technology Students' performance: A case of simple repeated measures approach. *Educ Inf Technol* 29, 18287–18297. <https://doi.org/10.1007/s10639-024-12593-3>
- [8] Hanlin Li, Yu Wang, Siqi Luo & Cui Huang. 2024. The influence of GenAI on the effectiveness of argumentative writing in higher education: evidence from a quasi-experimental study in China. *Journal of Asian Public Policy*, 1–26. <https://doi.org/10.1080/17516234.2024.2363128>
- [9] Kai Guo, Mengru Pan, Yuanke Li, Chun Lai. 2024. Effects of an AI-supported approach to peer feedback on university EFL students' feedback quality and writing ability. *The Internet and Higher Education*, 63, 100962. <https://doi.org/10.1016/j.ihed.2024.100962>
- [10] Chia-Yu Tsai, Yu-Ting Lin, and Ivan K. Brown. 2024. Impacts of ChatGPT-assisted writing for EFL English majors: Feasibility and challenges. *Educ. Inf. Technol.* 29, 22427–22445. <https://doi.org/10.1007/s10639-024-12722-y>.
- [11] Angela Choi Fung, Tam. 2024. Interacting with ChatGPT for internal feedback and factors affecting feedback quality. *Assess. Eval. High. Educ.* 50, 2, 219–235. <https://doi.org/10.1080/02602938.2024.2374485>.
- [12] Tao Zhou, Shuo Cao, Shujun Zhou, Yi Zhang, and Anming He. 2023. Chinese intermediate English learners outdid ChatGPT in deep cohesion: Evidence from English narrative writing. *System* 2303, 11812. 103141. <https://doi.org/10.1016/j.system.2023.103141>.
- [13] Kai Guo, Danling Li. 2024. Understanding EFL students' use of self-made AI chatbots as personalized writing assistance tools: A mixed methods study. *System*. 124, 103362. <https://doi.org/10.1016/j.system.2024.103362>.
- [14] Hui Li, Yawen Wang, Siyu Luo, and Chen Huang. 2024. The influence of GenAI on the effectiveness of argumentative writing in higher education: evidence from a quasi-experimental study in China. *J. Asian Public Policy*. 1–26. <https://doi.org/10.1080/17516234.2024.2363128>.
- [15] Yeonju Lee, Gwo-Jen Hwang, and Pei-Yi Chen. 2022. Impacts of an AI-based chatbot on college students' after-class review, academic performance, self-efficacy, learning attitude, and motivation. *Educ. Technol. Res. Dev.* 70, 5, 1843–1865. <https://doi.org/10.1007/s11423-022-10142-8>.
- [16] Chen Song and Yining Song. 2023. Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Front. Psychol.* 14. <https://doi.org/10.3389/fpsyg.2023.1260843>.
- [17] Mina Ghafouri, Jaleh Hassaskhah, and Amir Mahdavi-Zafarghandi. 2024. From virtual assistant to writing mentor: Exploring the impact of a ChatGPT-based writing instruction protocol on EFL teachers' self-efficacy and learners' writing skill. *Lang. Teach. Res.* 1, 23, 1–24. <https://doi.org/10.1177/13621688241239764>.
- [18] Mohammed Awad, Alafnan. 2025. DeepSeek Vs. ChatGPT: A Comparative Evaluation of AI Tools in Composition, Business Writing, and Communication Tasks. *J. Artif. Intell. Technol.* <https://doi.org/10.37965/jait.2025.0740>.