

DOI: <https://doi.org/10.63332/joph.v4i2.4188>

Evaluating AI-Generated Feedback in EFL Classrooms: Effects of Prompt Writing

Bandar Aljafen¹

Abstract

The introduction of Artificial Intelligence (AI) to English as a Foreign Language (EFL) classrooms has altered writing instruction through AI-mediated text generation and feedback. Given the difficulty of providing efficient personalized feedback to EFL learners' writing, this research explores the influence of prompts on the quality of AI-generated responses and their successive impact on learners' writing. The study concerns the effects of three types of prompts viz., corrective, analytical, and developmental given by 134 EFL learners enrolled for various English programs in the Department of English Language and Literature, College of Languages and Humanities, Qassim University, Saudi Arabia. The output generated by ChatGPT in response to these prompts was utilized by the sample for accomplishing writing tasks in English which were later analyzed to identify the most effective prompts. Moreover, a mixed survey with a combination of multiple choice, Likert-type, and open-ended questions to gain learners' perspectives on the efficacy of ChatGPT as an assistive writing tool was utilized to help triangulate findings. Descriptive statistics, ANOVA, paired t-tests, correlation, and Cronbach's Alpha were computed to make inferences. Findings showed positive attitudes to AI feedback ($M = 3.88-4.02$). ANOVA showed significant differences in the outputs generated by the three types of prompts ($F = 9.47, p < 0.001$), with the developmental prompt exhibiting the greatest improvement ($MD = 1.13$) in writing quality. The study concluded that prompt quality affects AI feedback and its efficacy in EFL writing, highlighting the significance of prompt engineering in EFL instruction.

Keywords: EFL writing, AI-mediated feedback, prompt engineering, ChatGPT, communicative competence, language learning pedagogy

Introduction

Writing is considered a challenging activity for EFL learners because it demands language knowledge as well as higher cognitive abilities with problems across language building blocks i.e., vocabulary, grammar, and syntax (Zhang & Hyland, 2023). However, Artificial Intelligence (AI) tools such as ChatGPT have changed educational settings with applications that enable individualized learning, computer-based feedback, and cutting-edge software loaded with unique features (Kohnke et al., 2023). ChatGPT comes with several benefits, such as facilitating the learning and teaching of English with instant feedback for syntax, mechanics, and vocabulary, and supporting global writing skills and language proficiency (Mohammed, 2023; Barrot, 2023). Moreover, accuracy and fluency can also be targeted with such AI applications that provide remedial feedback (Fitria, 2023).

AI systems enhance learning by identifying error patterns, proposing improvements, offering writing help, and minimizing performance anxiety through instant feedback that boosts motivation (Zou & Huang, 2023). Apart from accuracy and immediacy, the efficacy of feedback

¹ Department of English Language and Literature, College of Languages and Humanities, Qassim University, Buraydah, Saudi Arabia; Email: Jfn@qu.edu.sa



also depends on how it is perceived, judged, and used by learners, a process commonly known as response acceptance (Chen et al., 2024; Wang & Chen, 2024). Prompt-based AI applications such as ChatGPT provide instant, personalized feedback on vocabulary, grammar, and language skills. However, poorly designed prompts can result in irrelevant feedback and therefore no gains in skills. Hence, efficient learner-ChatGPT interaction facilitated by well-crafted prompts is vital to the effective use of ChatGPT in instruction (Bai & Wei, 2024)

Nevertheless, these applications raise many issues despite their advantages. Not every learner utilizes AI tools for personal development, and in environments where they are used, the risk of plagiarism is high as these programs can effortlessly produce sophisticated output whose originality is difficult to ensure (Pitychoutis, 2024). As AI tools grow in popularity, it is important to make sure they are used efficiently and fairly, particularly by Saudi EFL students whose educational environments are shaped by specific social, linguistic, and technical factors (Abduljawad, 2024). Besides AI-generated feedback, in terms of clarity, relevance, and utility, is influenced by the quality of prompts, whether developmental, analytical, or corrective.

Research Aim

This research aims to examine how various prompt types affect the quality of AI-generated feedback and how it affects the writing performance of EFL learners. In particular, it contrasts generic, grammar-focused, and EFL-specific inquiries to see how well they generate concise, pertinent, and helpful feedback for students' writing.

Related Works

Gains were reported in grammar, appropriacy of information, conciseness, and passive voice application in Ivenz's (2024) study with 110 undergraduates who used ChatGPT as a writing aid. In addition, qualitative data supported positive perceptions of the participants towards the use of the intervention. Metacognitive gains were reported in another study (Anjarani et al., 2024) that utilized ChatGPT with eighty Indonesian undergraduates. Though primarily used for idea generation, participants reported other gains such as global engagement with content as compared to the control group. Such findings make a case for including ChatGPT for cognitive and metacognitive gains in the language classroom.

In a case study with four language learners, Yan and Zhang (2024) examined the role of ChatGPT in feedback mechanisms and concluded that efficacy of usage was dependent on learners' proficiency both with language and technology though the participants reported negatively on the overall efficacy of ChatGPT in this study with some gains in engagement and interaction opportunities. Yatri et al. (2023) used three tools in a study with eighteen EFL learners and their teachers namely ChatGPT, Grammarly, and Elsa and reported positive perceptions on AI tools for speaking, writing, and translation but fear of losing on creative and critical thinking in learners. Teachers, on the other hand, indicated that AI can be a great learning companion but its reliability needs to be ensured with more research in diverse learning environments.

In a unique study, Baskara (2023) evaluated the efficacy of AI in developing EFL curricula and concluded that it has certain unique features that add to the quality of language classrooms such as personalization of content and pedagogy, and individual feedback which are harbingers of a general graduation towards its acceptability and integration into the learning process, in addition to the results indicating new trends in the use of these tools by teachers and learners and their attitudes to them. Wei et al. (2023) evaluated the role of Automated Writing Evaluation (AWE) with 190 Chinese EFL students in a repeated-measures randomized, controlled trial using Grammarly. Results indicated that task achievement, coherence, vocabulary, and grammar improved significantly with the application, but platform dependence and context-specific

sampling are some of the shortcomings of this study (Ahmed et al., 2023).

Ahmed (2023) applied factor analysis to interview responses to writing of 64 EFL learners in Saudi Arabia by comparing teacher-mediated and ChatGPT-assisted writing over four weeks. Greater fulfillment with instructor assistance was seen in the results, suggesting that while AI can improve knowledge, it cannot replace teachers. In addition to interviews, 181 Saudi EFL students participated in a survey in this study based on sociocultural theory and Technology Acceptance Model (TAM). Positive opinions and strong desire to utilize ChatGPT were evident in the results. Nevertheless, generalization of results is limited due to the small number of interviews.

Methodology

Participants in this study included 134 EFL students enrolled in the first semester of the year 2024 in different educational levels at the Department of English Language and Literature, College of Languages and Humanities, Qassim University, Saudi Arabia. The sample was mostly homogenous with similar academic backgrounds and intermediate English competence. Primary dataset comprised AI generated feedback to corrective, analytical, and developmental cues.

PARTICIPANTS

The study sample comprising 134 students in EFL was conversant with online writing tools and had already been exposed to online learning. The respondents represented diverse academic disciplines, though they shared similar language proficiency reflected in the University placement test. Demographic information of the sample is summarized in Table 1 below.

Table 1: Demographic information of study sample

Variable	Category	Frequency (n)	Percentage (%)
Gender	Male	62	46.3
	Female	72	53.7
Age Group	18–20 years	48	35.8
	21–23 years	56	41.8
	24–26 years	30	22.4
Academic Discipline	Humanities	38	28.4
	Social Sciences	34	25.4
	Business	29	21.6
	Science and Technology	33	24.6
English Proficiency Level	Intermediate	79	59.0
	Upper Intermediate	55	41.0
Experience with AI Tools	Less than 6 months	41	30.6
	6–12 months	52	38.8
	More than 1 year	41	30.6
Writing Performance Level	Moderate	63	47.0
	High	71	53.0
Feedback Clarity Perception	Agree	78	58.2
	Strongly Agree	56	41.8
Feedback Usefulness Perception	Agree	74	55.2
	Strongly Agree	60	44.8
Student Motivation	Moderate	59	44.0
	High	75	56.0

Writing Confidence	Moderate	61	45.5
	High	73	54.5

WRITING TASKS

The study sample was assigned an academic writing task based on the prescribed syllabus. These essays served as the input to produce AI-generated feedback based on multiple types of prompts to compare and evaluate the relationship of prompts to feedback and thence to writing quality.

PROMPT DESIGN

The prompts were of three types: developmental, analytical, and corrective. The following discussions break down the response generation process.

- **Corrective Prompt:** The application is asked to detect grammatical, lexical, and syntactical errors in the essays, identify mistakes, suggest corrections, and justify the corrections briefly.
- **Analytical Prompt:** The application is asked to evaluate the overall structure, consistency, readability, and coherence of the essays and highlight the strengths and weaknesses in learners' output.
- **Developmental Prompt:** The application is asked to suggest changes to enhance the quality of the essays with suggestions for concept development, clarity of argument, and diversity of word use, as well as general coherence.

AI FEEDBACK GENERATION

All the essays were graded upon the three types of prompts: corrective, analytical, and developmental.

Questionnaire Responses

Likert-scale items in the survey aimed to elicit the learners' perceptions about the feedback generated by ChatGPT using a five-point scale. The questionnaire measured five factors summarized in Table 2 below with some examples.

Table 2: Factors examined in the Questionnaire

Item Number	Aspect Measured	Example Statement
1	Feedback Clarity	"The AI-generated feedback was easy to understand."
2	Feedback Usefulness	"The feedback helped me improve my writing."
3	Writing Improvement	"Receiving AI feedback motivated me to revise my work."
4	Student Motivation	"The feedback increased my confidence in writing."
5	Writing Confidence	"Overall, I am satisfied with the AI-generated feedback."

STUDENT PERCEPTION SURVEY - FEEDBACK EVALUATION

Multiple-choice questions were used to determine which prompt style was considered most useful by the learners. These responses are summarized in table 3 below.

Table 3: Participants' responses to most useful prompt type

Survey Component	Objective	Example Items / Questions	Response Type
Multiple-Choice Comparisons	Identify which prompt type students found most effective	1. Which prompt provided the clearest feedback 2. Which prompt gave the most useful suggestions 3. Which prompt best supported overall writing improvement?	Corrective / Analytical / Developmental Prompt
Open-Ended Reflective Responses	Capture detailed student reflections on AI feedback	1. "Describe your experience using AI-generated feedback." 2. "What aspects of the feedback helped you improve your writing? What challenges did you face?" 3. "How could the AI feedback be improved?"	Open-Ended Text Responses

DATA ANALYSIS

Survey data was analyzed using SPSS version 28 through descriptive statistics, one-way ANOVA, paired sample t-tests, Pearson's correlation, and Cronbach's Alpha.

Descriptive statistics (DS): DS summarizes the important characteristics of information, such as performance of the learners, types of prompts, and quality of feedback provided by ChatGPT to identify general patterns, central tendencies, and variability represented in Equation (1).

$$\bar{Y} = \frac{\sum_{j=1}^o Y_j}{o} \quad (1)$$

Where Y_j is the j -th learners' experiential mark, and o is the total number of participants. The study compared output across prompt types and evaluated the effect of automatic feedback on EFL learners using mean \bar{Y} . Equation (2) illustrates how the standard deviation was calculated to measure the extent or diversity of learners' results around the mean.

$$t = \sqrt{\frac{\sum_{j=1}^o (Y_j - \bar{Y})^2}{o-1}} \quad (2)$$

Where t represents the outcome of the j -the data point, and j is the ranking for each observation from 1 to o .

σ is the computed test statistic (standard deviation). Standard deviations and means help analyze learner development consistency and AI input efficiency; smaller variances suggest score attention, while bigger deviations show better variation.

One-way ANOVA: One-way ANOVA is a statistical method of comparing the means of a number of independent groups of samples in which samples are selected at random and population distributions are assumed to be normal with equal variances, demonstrated in Equation (3).

$$F = \frac{(o-1)SS_c}{[(s-1)SS_x]} \quad (3)$$

Where SS_c is the entire of squared values within categories (capturing unpredictability within each prompt type), SS_x is the total number of squares for each group (catching shifts owing to

variations among prompt types), s is the sum of prompt kinds, and O is the total number of instances across all groups.

Paired t-test: The paired t-test helped determine the extent to which the AI feedback causes significant increase in writing quality by comparing results of the same group of participants prior to and after the intervention. The test statistic (t) is computed in Equation (4).

$$t = \frac{\sum e}{\sqrt{\frac{o(\sum e^2) - (\sum d)^2}{o-1}}} \quad (4)$$

Where $\sum e$ shows the difference between paired observations, o is the entire quantity of paired explanations and also indicates the degrees of freedom, $\sum d$ is the sum of all paired differences, and $\sum e^2$ is the sum of squared paired differences. This method enabled the research to statistically ascertain whether learners' writing skills significantly improved as a result of AI feedback.

Cronbach's Alpha: Internal reliability of survey questions was assessed using Cronbach's Alpha represented in Equation (5).

$$\alpha = \frac{o\bar{d}}{w + (o-1)\bar{d}} \quad (5)$$

Where O is the quantity of substances, \bar{d} is the mean variance between components, and \bar{w} is the mean variability of individual items. The validity of the research's conclusions about the efficacy of AI feedback is ensured by a higher alpha value, which shows that the instrument consistently captures learners' perspectives and involvement.

I. Result

This study aimed to investigate the efficacy of AI-generated feedback in EFL writing by examining learners' output and perceptions. Results show trends, connections, and notable variations in how AI-mediated feedback helps learners improve their writing as discussed below.

A. DESCRIPTIVE STATISTICS

Table 4 and Figure 3 demonstrate the DS of the opinions of students regarding AI-generated feedback in EFL writing tasks. Grammar and sentence improvement reflect the maximum mean score ($M = 4.02$, $SD = 0.69$). Strong learner agreement is indicated by high mean score ($M = 3.97$, $SD = 0.68$) for academic writing help. While concept development shows clear understanding ($M = 3.91$, $SD = 0.70$), essay organization also receives a favorable assessment ($M = 3.94$, $SD = 0.72$). Furthermore, learners' motivation to edit their writing was positively impacted by AI-assisted feedback ($M = 3.88$, $SD = 0.75$).

Table 4: DS of responses on AI-generated writing support (N=134)

Items	Feedbacks	N	Mean	Std. Deviation
1	The AI feedback aided me recover my grammar and sentence structure.	134	4.02	0.69
2	The suggestions provided by AI helped me organize my essay better.	134	3.94	0.72
3	The feedback encouraged me to revise and improve my writing.	134	3.88	0.75
4	The AI explanations helped me understand how to develop my ideas clearly.	134	3.91	0.70
5	The feedback supported my learning process in academic writing.	134	3.97	0.68

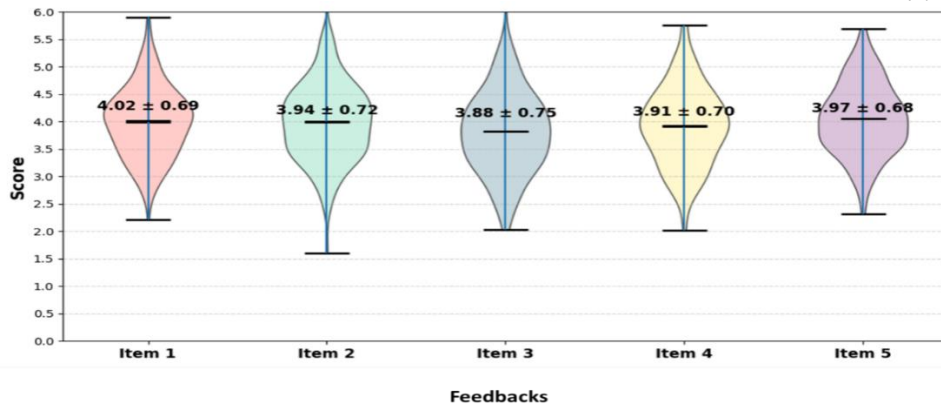


Figure 3: Student perceptions of AI feedback in EFL writing
RELIABILITY ASSESSMENT OF AI FEEDBACK PROMPT TYPES

Table 5 and Figure 4 show the reliability analysis of survey items that were associated with different types of AI feedback prompts. It is evident that developmental prompt has the highest internal consistency (Cronbach’s Alpha = 0.90), followed by analytical prompt ($\alpha=0.88$), and corrective prompt ($\alpha =0.86$). Each type of prompt had five survey questions that examined learners’ perceptions regarding AI-generated feedback. The overall reliability value of 0.89 of all the items used as a measurement tool in the research is confirmation of the reliability of the tool.

Table 5: Internal consistency of survey items

Prompt Type	Number of Items	Cronbach’s Alpha
Corrective Prompt	5	0.86
Analytical Prompt	5	0.88
Developmental Prompt	5	0.90
Overall Reliability	15	0.89

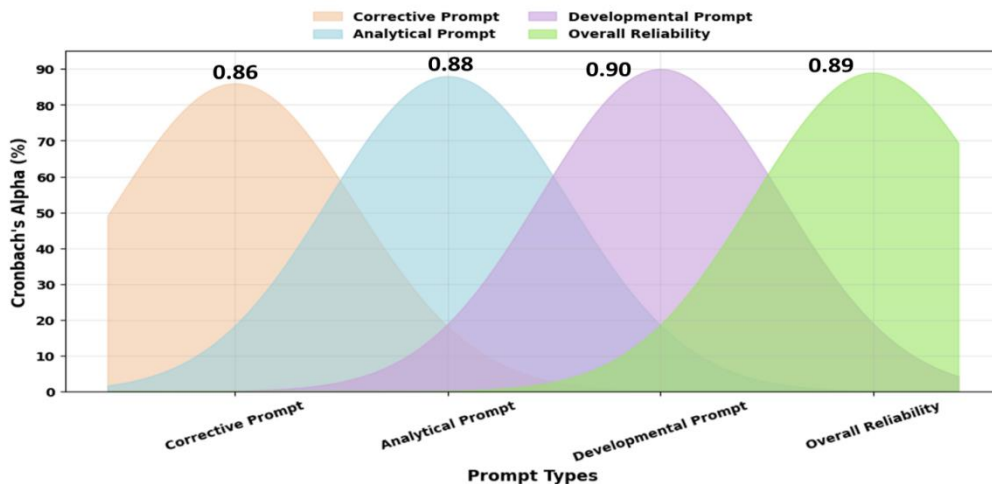


Figure 4: Cronbach’s Alpha showing high reliability of corrective, analytical, and developmental AI feedback prompts

COMPARATIVE ANALYSIS OF PROMPT TYPES USING ANOVA

Results of one-way ANOVA are represented in Table 6. The total squares among collections with two grades of liberty are 345.62, with a mean square of 172.81. The within-group variance mean square is 18.08, and the sum of squares is 2367.45, with 131 degree of freedom. The F value of 9.47, with a significance level of $p < 0.001$, shows that the differences between the types of prompts are statistically significant.

Table 6: One-way ANOVA results

Source	Sum of Squares	df	Mean Square	F	p
Between Groups	345.62	2	172.81	9.47	< 0.001
Within Groups	2367.45	131	18.08		
Total	2713.07	133			

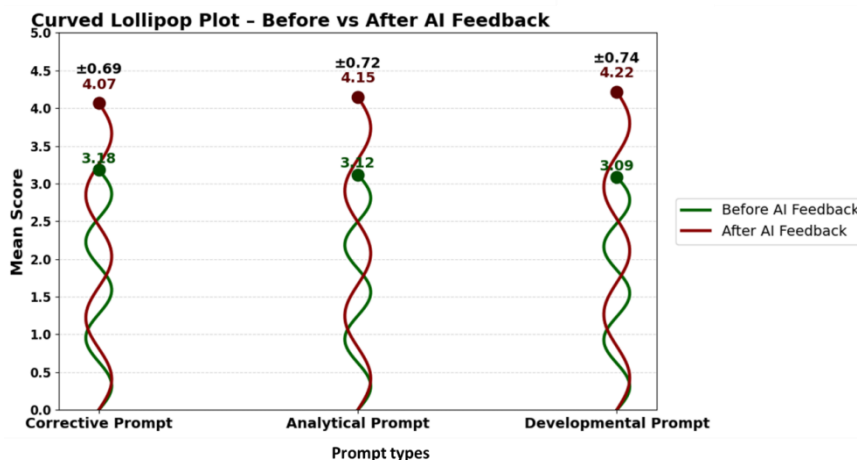
(Note: df = Degrees of Freedom; F = F-statistic; p = significance level)

EFFECT OF AI FEEDBACK PROMPTS ON SCRIPT

Table 7 and Figure 5 show that following AI-assisted feedback based on all three prompt kinds, notable improvements could be noted in the writing output. While analytical prompts improved writing means from 3.12 to 4.15 (MD = 1.03, SD = 0.72, $t = 11.08$), corrective prompts indicated writing improvement from $M = 3.18$ to $M = 4.07$ (MD = 0.89, SD = 0.69, $t = 10.62$). The biggest effect was noted as a result of developmental prompts which increased the writing means 3.09 to 4.22 (MD = 1.13, SD = 0.74, $t = 11.74$).

Table 7: Paired sample t-test

Prompt Type	Mean (Before AI Feedback)	Mean (After AI Feedback)	Mean Difference	Std. Deviation	t value	df	p value
Corrective Prompt	3.18	4.07	0.89	0.69	10.62	133	< 0.001
Analytical Prompt	3.12	4.15	1.03	0.72	11.08	133	< 0.001
Developmental Prompt	3.09	4.22	1.13	0.74	11.74	133	< 0.001

**Figure 5:** Outcome of paired sample t-test

CORRELATION ANALYSIS AMONG AI FEEDBACK PROMPT TYPES

Table 8 and Figure 6 demonstrate Pearson's correlation analysis of the three types of AI feedback prompts showing high positive correlation between corrective and analytical prompts ($r = 0.69$). Also, developmental and corrective prompts are strongly correlated ($r = 0.72$). The strongest association ($r = 0.75$) is between analytical and developmental stimuli.

Table 8: Pearson's correlation

Variables	Corrective Prompt	Analytical Prompt	Developmental Prompt
Corrective Prompt	1	0.69	0.72
Analytical Prompt	0.69	1	0.75
Developmental Prompt	0.72	0.75	1

Note: Correlation is significant at $p < 0.01$

Figure 6: Pearson correlation matrix showing the relationships among corrective, analytical, and developmental AI feedback prompt types

STUDENT RESPONSES TO DIFFERENT AI FEEDBACK PROMPTS

Analysis of learners' feedback on AI shows their highly motivated state to make changes, 63 percent reported that they wrote better post feedback, and 53 percent stated that they comprehended the information better as a result of the feedback. Prompts that focused on grammar proved to be moderately effective, and general prompts were less effective, especially in the development of writing, shown in Table 9.

Table 9: Percentage distribution of learners' rating of AI prompts

Prompt Type	Encourages Essay Revision (%)	Enhances Writing Development (%)	Provides Clear Explanations (%)
Corrective Prompt	69	63	53
Analytical Prompt	19	38	41
Developmental Prompt	13	0	6

Discussion

Findings in this research show that prompt design has a major impact on the efficacy of feedback, with developmental prompts offering the most helpful feedback for enhancing learners' writing. This research builds on EFL writing pedagogy by indicating that the quality, clarity, and use of AI-generated feedback are significantly enhanced by strategically designed prompts; hence, the weaknesses of generic AI support are overcome. Focusing on the interaction between learners and the fast-engineering process in AI bridges the existing deficits in AI-mediated writing systems and facilitates independence, improved writing, and communication abilities.

Conclusion

The use of AI tools such as ChatGPT in EFL classrooms has transformed writing instruction due to the rapid feedback provided by the tool. However, to achieve the maximum effect, it is imperative to understand the role of prompt design in relation to writing performance, interest, and quality of feedback. By examining 134 Saudi EFL learners' writing and the impact of AI feedback on it to determine effective methods to use artificial intelligence tools in language training, this research aims to improve the quality of learning in the EFL classroom. The study

concluded that AI-generated feedback using corrective, analytical, and developmental prompts is effective in improving EFL learners' writing. The data was collected using a Likert-scale based survey and data was analyzed using DS, one-way ANOVA, paired t-tests, and Cronbach's Alpha (using SPSS). Findings demonstrate good dependability ($\alpha = 0.86\text{--}0.90$) and favorable opinions of AI-generated feedback ($M = 3.88\text{--}4.02$). Significant differences were found by ANOVA ($F = 9.47, p < 0.001$). The developmental prompt showed the biggest gain in performance ($MD = 1.13, t = 11.74$).

Recommendations

Practical Implications: A combination of thoughtfully crafted prompts and AI-based feedback improves EFL writing instruction, fosters learner autonomy, and raises writing quality. Instructors can be advised to combine technology and pedagogy most effectively, which cannot be achieved without the current classroom results.

Social Implications: The research focuses on the responsible use of AI in education, stimulating a critical attitude towards the use of technology in EFL learning communities, promoting equitable access to individualized learning, and increasing collaborative skills.

Limitations and Future Directions

Using a smaller number of participants from one University and relying on comments written by ChatGPT represent limitations for this study. Larger sample sizes, diverse institutions, and the use of multiple AI applications for comparison would provide additional information on the long-term effects of technology on EFL writing in general.

Acknowledgement:

The Researchers would like to thank the Deanship of Graduate Studies and Scientific Research at Qassim University for financial support (QU-APC-2024).

References

- Abduljawad, S. (2024). Examining the impact of artificial intelligence (AI) tools on Saudi Arabian ESL students' writing skills. *Erudita: Journal of English Language Teaching*, 4(2), 102–115.
- Ahmed, H., Kausar, S., & Mohammed, O. (2023). Social networking sites mediated course reflections on instructional technology: A case study of female students at women's university. *TESOL and Technology Studies*, 4(2), 36–48. <https://doi.org/10.48185/tts.v4i2.836>
- Ahmed, M. A. (2023). ChatGPT and the EFL classroom: Supplement or substitute in Saudi Arabia's eastern region. *Information Sciences Letters*, 12(7), 2727–2734. <https://doi.org/10.18576/isl/120704>
- Anjarani, S., Furqon, M., & Rakhmawati, F. (2024). ChatGPT to foster students' engagement in writing class: An intervention study. *Voices of English Language Education Society*, 8(3).
- Bai, L., & Wei, Y. (2024). Exploring EFL learners' integration and perceptions of ChatGPT's text revisions: A three-stage writing task study. *IEEE Transactions on Learning Technologies*, 17, 2161–2172.
- Barrot, J. S. (2023). Using ChatGPT for second language writing: A systematic review. *Computers and Composition*, 69, Article 102789.
- Baskara, F. R. (2023, November). Revolutionising EFL curriculum: A theoretical analysis of generative AI for active learning. In *Proceedings of the UNNES-TEFLIN National Conference* (Vol. 5, pp. 278–293).

- Chen, X., Xie, H., & Hwang, G. J. (2024). A multi-perspective study on AI-generated feedback acceptance in EFL writing classrooms. *Education and Information Technologies*, 29(3), 3451–3472.
- Fitria, T. N. (2023). Artificial intelligence (AI) technology in OpenAI ChatGPT application: A review of ChatGPT in writing English essay. *ELT Forum: Journal of English Language Teaching*, 12(1), 44–58.
- Huang, J. (2023). Engineering ChatGPT prompts for EFL writing classes. *International Journal of TESOL Studies*, 5(4), 73–79. <https://doi.org/10.58304/ijts.20230405>
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, 54(2), 537–550.
- Mohammed, O. S. M. (2023). Yemeni students' utilization of Google Translate for learning English language: Use, advantages, and disadvantages. *Journal of Social Studies*, 29(4), 136–156.
- Pitychoutis, K. (2024). Harnessing AI chatbots for EFL essay writing: A paradigm shift in language pedagogy. *Arab World English Journal (AWEJ), Special Issue on ChatGPT*.
- Polakova, P., & Ivenz, P. (2024). The impact of ChatGPT feedback on the development of EFL students' writing skills. *Cogent Education*, 11(1). <https://doi.org/10.1080/2331186X.2024.2410101>
- Wang, Y., & Chen, L. (2024). EFL learners' perceptions of AI-generated corrective feedback: A mixed-methods study. *Language Learning & Technology*, 28(1), 1–22.
- 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>
- Yan, D., & Zhang, S. (2024). L2 writer engagement with automated written corrective feedback provided by ChatGPT: A mixed-method multiple case study. *Humanities and Social Sciences Communications*, 11(1), 1–14.
- Yatri, D., Anugerahwati, M., & Setyowati, L. (2023). Artificial intelligence (AI) in language learning (English and Arabic class): Students' and teachers' experience and perceptions. *Transformational Language, Literature and Technology Overview in Learning (Transtool)*, 3(1), 1–12.
- Zhang, Z., & Hyland, K. (2023). Student engagement with teacher and automated feedback in L2 writing. *Journal of Second Language Writing*, 61, Article 101034.
- Zitouni, M., Al-Traif, H., Zemni, B., Mohammed, O. S., & Aljasser, M. (2021). Utilization of YouTube to improve the pronunciation skill of Saudi learners in translation departments. *Asian ESP Journal*, 17(2), 133–154.
- Zou, D., & Huang, X. (2023). The effects of AI-powered feedback on EFL students' writing motivation and performance. *Interactive Learning Environments*, 31(8), 5123–5138.