Is ChatGPT a Cure all? Demystifying the Impact of Using ChatGPT on EFL Learners' Writing Skill

Parichehr Fereidouni[1] iD
Majid Farahian[2*] iD

[1]M.A. Student in Applied Linguistics, Department of ELT, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran
[*2]Assistant Professor of Applied Linguistics, Department of ELT, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran

ABSTRACT

The aim of the current study was to demystify the impact of using ChatGPT on English as a foreign language (EFL) learners' writing skill. To this end, 45 EFL male and female students from two language schools participated in this study, selected via convenience sampling from 67 learners based on a placement test. The participants were divided into three groups. In the first experimental group the students practiced writing using ChatGPT, while in the second experimental group, they received writing instructions using ChatGPT and teacher input, and in the third, the participants followed the conventional method was utilized. Before and after the treatment, the participants were subjected to pre and posttest. Finally, the performance of the three groups was compared via one-way ANOVA. In addition, 10 EFL learners from both experimental groups were selected voluntarily to take part in the interview and the qualitative data were analyzed based on the theme-based analysis. The findings revealed that the second experimental group outperformed the first experimental and the control groups, revealing the efficiency of the instruction via ChatGPT along with teacher input. The result of the interview showed that while the EFL students generally hold a favorable view of ChatGPT, they expressed concerns about the use of the artificial intelligence (AI) tool. In addition, they lack the necessary skills to effectively employ it to help them write. This deficiency could be a significant obstacle to the implementation of technology in their EFL writing.

KEYWORDS: Artificial intelligence; EFL learners; CALL; ChatGPT; Writing skill

1. Introduction

An increasing amount of research has acknowledged the significance of writing skills for EFL students (Hasanah & Fatimah, 2020; Rovikasari & Supriyadi, 2020). Writing is a language skill that is intricately connected with the other facets of a language (Fitria, 2021). An EFL writer must be well-versed in selecting the right structure, language, writing mechanics, vocabulary, and more prior to commencing the writing process (Lutai & Besarab, 2020). On top of that, thoughts, words, and sentences in a piece of writing come from the author's essential abilities including his own creativity uniqueness, cognitive, and metacognitive abilities (Farahian & Ebadi, 2023). As such EFL writing has turned out to be a demanding task for EFL learners (Nguyen & Suwannabubpha, 2021) and one of the challenges is inaccurate grammar (Ahmed, 2019); therefore, they need assistance in the writing process.

In recent decades, there has been a great deal of propensity toward using computers for language teaching and learning. However, according to Beatty (2003), with the advent of technological advancement in the multimedia era especially computing

and the Internet, the significant status of computers in language instruction has turned out to be undeniable. Regarding English language teaching (ELT), several high-end software packages have been employed to meet the needs of EFL learners (Jahanban-Isfahlan, et al., 2017). In fact, technology has significantly changed the way people learn. Thanks to technology, students may now access a wealth of material online. With this access, students may no longer depend only on textbooks, increasing the accessibility and inclusivity of information. AI tools are becoming more and more popular these days over the past few decades, this technology has grown, better yet, nowadays, a wide variety of robots or AI in all shapes and sizes are used in various programs and applications. Among the many AI technologies that are available, ChatGPT is one of the most widely used ones that may be employed in writing classes. It has been said that ChatGPT is a simple application that can assist students by reviewing their writing for spelling grammar and punctuation as well as writing mistakes (Khazaie et al, 2013).

Additionally, ChatGPT offers thorough and helpful feedback with corrections and suggestions to improve the writing's clarity, precision, effectiveness, readability, and impact with a high evaluation speed and accuracy rate (Grammarly, 2020). Using AI for error correction can be a very effective way to provide EFL students with personalized language, immediate feedback, and engaging and interactive learning experience (Polakova & Ivenz ,2014; Slamet, 2024). They are available 24/7 and can be accessed from anywhere with an internet connection, making them a convenient tool for learners who may not have access to traditional learning resources. AI tools have the potential to enhance learning experience for students and provide a valuable tool for educators looking to innovate and improve their teaching methods. In general, feedback plays a crucial part in the process of teaching and learning English to EFL students since it enhances students' skills and boosts their motivation (Cao et al., 2022).

The incorporation of ChatGPT in foreign language education (FLE) has been a significant focus of recent research, as evidenced by several studies (Hakiki et al., 2023; Jagdishbhai & Thakkar, 2023; Katar et al., 2023; Li, 2024). ChatGPT serves as an effective tool for FLE, allowing students to engage with it based on their English proficiency level and receive immediate feedback on grammar, punctuation, vocabulary, and more. Additionally, ChatGPT aids in text comprehension, offering summarization or simplification and providing explanations, examples, additional learning resources, exercises, and more. It also functions as a translation tool, capable of translating to or from English (Wang et al., 2020). On the other hand, the literature (e.g., Moghadam, 2023) shows that in the era of AI's influence and expansion, the university education system in Iran is at a crossroads, with the best-case scenario being evolution and the worst-case scenario being collapse. The actions and intentions of stakeholders such as policymakers, administrators, faculty, students, researchers, and non-academic employees will determine which scenario becomes a reality. Currently, there is no national or university plan in place to take advantage of AI's capabilities and address its potential threats. As a result, the Iranian university system is currently focused on maintaining the status quo, which may not be sustainable without a plan to transition to a more desirable situation. In addition, there is a scarcity of research especially in the EFL context exploring the possible uses of GenAI in settings like Iran, which faces numerous challenges in its education system (Tafazoli, 2024). This highlights a significant gap in our understanding of how to use new technologies to address important educational needs. Most current studies concentrate on incorporating GenAI in settings with ample infrastructure, technology access, and educational resources. Consequently, this study seeks to address existing gaps in the literature by examining the impact of ChatGPT on the writing skills of EFL learners and gathering EFL students' perceptions of ChatGPT's potential for enhancing their writing. This investigation seeks to reveal distinctive insights and practices that could have wider implications for language education across the globe. By exploring how ChatGPT affects EFL writing abilities and their perceived worth, this research can provide useful lessons and strategies that could be tailored and implemented in various educational contexts worldwide. Moreover, by evaluating the impact of ChatGPT on language learning experiences, this study intends to enhance the ongoing conversation about the role of AI technologies in influencing the future of EFL writing skill.

2. Literature review

Integrating computer technology and Internet-based syllabi in teaching may be of great importance. According to Hayati (2005), the benefits of using computers in language learning include improving grammar, structure, and intonation, giving learners freedom, testing learners' knowledge, self-evaluation, and more contact with language in different situations.

Computer-assisted language learning (CALL) began in the 1950s and 1960s, mostly in the USA. Pioneers such as Suppes, Kemeny/Kurtz, and Bitzer were among the first to utilize a computer as part of the learning process. Bitzer was one of the first to recognize the significance of graphics and sound in the teaching process. As such, CALL was designed to increase learners' success and teachers' quality at all levels. Recently, development has been fast and technology has been considered as a supplementary teaching material.

As noted by Meskill and Mossop (1997), computers promote learning by creating an engaging environment and boosting motivation. They can assist shy learners who may fear making errors in a traditional classroom. Computers serve as valuable online resources, which is advantageous in language acquisition and can accommodate learners with varying skill levels. Furthermore, the capacity to provide immediate feedback is another advantage of CALL. However, CALL is not without its shortcomings. Dizon and Gayed (2021) suggest that self-access programs can sometimes leave learners feeling excessively

independent, potentially leading to feelings of being overwhelmed by the vast amount of information and resources available. Alternatively, if classroom techniques are simply translated to an online format, students may experience an overload of guidance from the computer. It is important to avoid relying on flashy multimedia techniques, while still considering relevant theories on language acquisition. However, this does not imply that multimedia should be disregarded altogether. Some researchers, such as Meskill and Mossop (1997), argue that effective multimedia applications can enhance learning outcomes. With such a perspective AI intelligence was widely used in EFL learning and teaching.

With the explosion of data resources, advancements in computing power, and generational shifts in data transmission, AI has entered its third wave since Jacobsen and Lock (2005) introduced the concept of "intelligence" in the 1950s. This comprehensive and rapid development of AI brings forth numerous possibilities and approaches that have transformed the world. One of the most noticeable aspects of this transformation is the implementation of AI-based technologies such as machine vision, data mining, text analysis, and speech recognition in real-world applications across various settings (Dwivedi, et al., 2023). Consequently, the evolution of machine learning, neural networks, and deep learning is occurring within the realm of language learning, similar to its impact in other domains (Patty, 2024).

Education benefits from AI technology, which analyzes individual student data to offer personalized instruction, customizing feedback, and learning materials according to each student's unique needs and preferences. This tailored approach is particularly beneficial for special education students who need individualized teaching (Alenizi et al., 2023). Furthermore, AI technology can deliver instant feedback and assistance that provide immediate responses to student assignments and encourage interactive learning experiences. This technology helps teachers save time, allowing them to focus on more advanced instructional tasks (Chen et al., 2022).

One of these emerging technologies is ChatGPT. ChatGPT refers to an advanced chatbot that uses OpenAI's GPT-3 technology. It has been improved with different types of training methods. Reinforcement learning is a method where the machine learns by trying things out in its environment to get the most rewards while exploring different options on its own (Ram & Verma, 2023). ChatGPT can talk with users naturally and fluently. The language model uses a network that learns from a lot of data to create different connections. This helps ChatGPT generate text that sounds like human language. It can answer follow-up questions, admit when it is wrong, correct false claims, and turn down inappropriate requests (Muller & Gregoric, 2017). ChatGPT can also create different types of writing, essays, jokes, and poems. With constant feedback from users, ChatGPT can get better at doing similar tasks. ChatGPT cannot 'understand' the text it creates or the meaning behind the information. This often leads to answers that sound reasonable but are actually wrong or make no sense. It seems like you might have entered a reference or a code. Researchers have raised worries about how accurate the information from ChatGPT is (Graham, 2022). There have been instances where ChatGPT created made-up content when it did not have enough information and even invented fake sources (Rudolph et al., 2022). Experts noted that even though AI tools seem good at putting words together, they actually just learn how words are related to each other instead of really understanding what the words mean. In simple terms, ChatGPT is a modern tool that creates text and does not browse the Internet for current information.

In a study, Al-Jarf (2004) explored how online learning and traditional learning affected the writing skills of EFL students. He discovered that using online lessons alongside regular classes greatly helps improve writing structure. The study looked at how teaching tools and online learning affect education. The results showed that there is an important link between how well students learn and online teaching. The experimental group outperformed the control group.

Rahimi and Yadollahi (2011) studied perceptions of learning English with computers that improved their writing skill. A survey was used to find out how students perceive learning languages with computers. The study found that Iranian female students had a positive opinion about using computers for learning.

In another study involving Turkish university students learning and their use of AI tools, Han and Sari (2022) found that the group that was provided with both automated feedback and teacher feedback improved more than the group that only received feedback from the teacher. Likewise, in a study by Dizon and Gayed (2021), Japanese EFL students who used the mobile version of Grammarly saw improvements in their grammar and vocabulary. However, their sentence structure and speaking flow did not change much.

Wilkens (2020) studied two Chinese university students who used AI for four weeks. The study found that the students appreciated getting feedback that pointed out mistakes in their first language. They were able to use this feedback to find their own errors. Based on the results people who create AI should improve their resources, offer more choices, and give the option to use L1 or just L2. Harvey-Scholes (2018) also supported the idea that allowing students to correct their own work is effective since it helps them learn even when the teacher is not around.

Jiang and Yu (2022) studied how a group of EFL students experienced automated feedback on their writing. They highlighted the importance of helping students understand how to use the feedback and resources effectively. They discovered that specific error advice was more helpful than general feedback.

Wambsganss et al. (2022) investigated the use of AI for writing that offers automatic feedback to EFL learners. They did not see much difference in the writing improvements between the group that got automated feedback and the group that got regular feedback. However, the group that received prompts to compare themselves to others produced better quality writing with stronger arguments.

Concerning the utilization of AI in improving language skills, previous studies (Abdullayeva & Musayeva, 2023; Ausat et al., 2023; Johnson et al., 2019; Shidiq, 2023; Xue & Wang, 2022; Yusriadi et al., 2023) confirmed the efficacy of ChatGPT as an AI tool on language skills. However, there is currently a dearth of empirical data regarding ChatGPT's effect on students' writing skill (Su et al. in 2023). In addition, the teacher feedback along with ChatGPT, as well as the opinion of EFL students regarding the use of ChatGPT for error correction has not been investigated in a single study yet. Therefore, the present study examined the impact of using ChatGPT, an AI-powered chatbot, on the writing skills of EFL learners, and investigated the perception of Iranian EFL students about the implementation of ChatGPT for writing error correction. As such, the following research questions were proposed in line with the purpose of the study:

RQ1. Compared to the control group (the conventional teacher feedback) and the first experimental group (teacher feedback plus using ChatGPT), does the implementation of ChatGPT alone improve EFL students` performance in writing in English?

RQ2. What is the perception of EFL students of using ChatGPT to correct their errors?

3. Methodology

3.1. Design

To answer the research questions, the current study used a mixed design to profit from the two integrations. The following figure presents the design of the study.
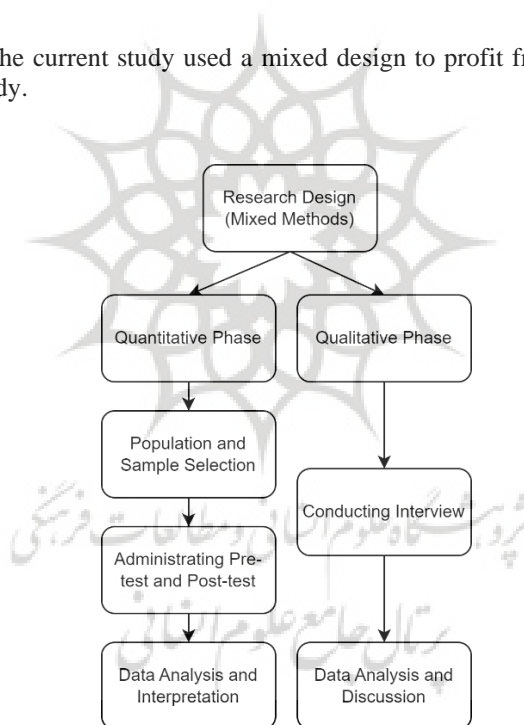

Figure 1. The design of the study

3.2. Participants

To perform the study, 45 EFL male and female students from two reputable language schools in Kermanshah, Iran participated in this study, selected via convenience sampling from 67 learners based on a placement test. As part of their syllabus, the students were taught the elements of paragraph writing and some basic types of paragraphs such as narrative. The age range of the participants was 16 to 28. Due to the nature of the study, which necessitated the participants' familiarity with technological tools, it was decided to select the participants from intermediate levels. The participants were divided into three groups. In the first experimental group they practiced using ChatGPT, while in the second experimental group, they received writing instructions using ChatGPT and the teacher's input, and in the third, participants followed the teacher's instructions. The participants were assured of the confidentiality of their responses and that their identities would remain anonymous. It should

be noted that 10 participants participated in the interview.

3.3. Instruments

3.3.1. Oxford placement test

To assess the English language proficiency of the participants, the current study utilized a general training version of the Oxford Placement Test (OPT). The OPT is recognized as a valid and reliable assessment tool and serves as an efficient means of categorizing participants, which is utilized in the initial stage of participant selection for English language studies (Zolfaghari, 2023).

3.3.2. Writing pre-test and post-test

To assess the extent of plagiarism and overlapping content among master's theses, a corpus comprising 43 Translation theses and 46 TEFL (Teaching English as a Foreign Language) theses were downloaded from the Irandoc database and subjected to analysis using plagiarism detection software. Specifically, the iThenticate tool was utilized to quantify the degree of plagiarized material present within these academic works.

3.3.3. Interview

For the qualitative aspect of the study, a semi-structured interview was administered because it allowed the participants to be heard in their own voices and open-ended questioning helped the researcher gain a richer understanding of participants' perceptions and knowledge. The interview has five items. It should be noted here that to ensure the validity, the items were reviewed by three language experts (Appendix B).

3.4. Procedure

To conduct the study, first, 67 male and female Iranian EFL students from a language institute in Kermanshah, Iran, were chosen based on the convenience sampling. They then took the OPT. Regarding the OPT results, 45 EFL learners were selected, and divided into three equal groups. Before the treatments, the writing pretest was administered to three groups. In the first week, an instructor introduced the course structure, narrative writing, and course materials. Following a brief segment of a film, the students were tasked with writing a narrative paragraph which lasted for 15 minutes. From weeks 2 to 12, the students participated in narrative writing instruction and completed three narrative writing assignments in class under the supervision of the teacher.

As for the participants of the first, second and the third experimental groups, they also had one-hour training on using ChatGBT. Additionally, the students were taught various ChatGPT prompts to seek suggestions and make edits to enhance the clarity, grammar, and overall coherence of their written work. The treatment took twelve sessions. At the last session, the writing posttest was conducted.

More specifically, in the first experimental group, the participants received treatment using ChatGPT. In this group, ChatGPT was utilized as a tool to give feedback on their drafts including grammar, vocabulary, organization, coherence, sentence structure, and the overall coherence and organization of their writing. In addition, ChatGPT generated writing exercises that the students could complete to practice writing skill. Finally, ChatGPT was used to provide language input by recommending some extra paragraph writing exercises and notes.

Additionally, The second experimental group was instructed with ChatGPT along with teacher input. The only difference between this group and the first experimental group was that the participants not only received feedback from ChatGPT but were also provided by the teacher's feedback on their writing errors. In contrast, the control group was taught only via the teacher's instructions. In this group, participants were initially required to revise the texts based on the feedback they received from ChatGPT. If there were any ambiguity or points suggested by the AI tool they did not accept or understand they handed their drafts to the teacher to be checked.

The control group participated in regular writing sessions, completed designated writing assignments, and received feedback from their teacher. Unlike the experimental group, the control group did not benefit from AI-assisted insights from ChatGPT. They relied on feedback from the teacher, who offered guidance based on her teaching experience. Participants in the control group engaged in various writing exercises and activities that encompassed a combination of classroom tasks and topics of interest, similar to what the experimental group encountered. These activities aimed to address different aspects of writing, including grammar, vocabulary, organization, coherence, and sentence structure. The teacher provided personalized feedback on writing assignments, pointing out areas for improvement and suggesting ways to enhance their work. Feedback

for those in the control group was given solely during the teacher-led writing sessions, which were held regularly to ensure continuous support and input.

Finally, 10 EFL learners from both experimental groups were selected voluntarily to take part in the interview. Each interview lasted 15 minutes, conducted whether online via Skype app or face to face. The data obtained from the the interviews were then gathered and subjected for analysis.

3.5. Data analysis

The gathered data were analyzed both quantitatively and qualitatively. In fact, after encoding the data collected through writing pretest and posttest scores as well as the interview, the data were entered into SPSS version 22 for data analysis. Based on the research questions, for the quantitative data, in order to present descriptive statistics, means, minimums, maximums, and standard deviations of each variable and the corresponding sub-factors were computed. To this end, the performance of the three groups was compared via one-way ANOVA. The qualitative data were analyzed based on the theme-based analysis.

4. Results

The assumptions underlying the use of parametric statistical tests like ANOVA include the normality of distribution, the homogeneity of variances, the requirement for at least interval-level variables, and the independence of measurements. To evaluate the first assumption, we calculated the kurtosis and skewness values along with their respective z-scores for the groups.

Concerning the writing abilities of the groups, the skewness z-score was Z skewness = -0.116, while the kurtosis z-score was Z kurtosis = 0.722, suggesting that a value exceeding 1.80 is significant at $p < .05$. Therefore, none of the z-scores surpassed 1.80, indicating that the scores follow a normal distribution. Additionally, Levene's test was conducted to assess the second assumption, which pertains to the homogeneity of variances. The results are presented in Table 1.

Table 1. Levene's test results

	Levene Statistic	df1	df2	Sig.
Based on Mean	.081	1	44	.732
Based on Median	.072	1	44	.722
Based on Median and with adjusted df	.073	1	43.02	.734
Based on trimmed mean	.084	1	45	.728

Based on Table 1, Levene's test indicates a non-significant outcome at $p \leq .05$. Therefore, it can be inferred that the variances across the groups are not significantly different and are roughly equal, indicating that the assumption of homogeneity of variances is met. Since the key assumptions for the parametric test are satisfied, a one-way ANOVA—being a parametric test—can be performed.

As stated, to have a homogeneous sample of the participants, OPT was run and analyzed. The descriptive statistics of the OPT results are illustrated in Table 2.

Table 2. Descriptive statistics of the OPT

	N	Min	Max	M	SD
OPT	67	44	56	47	1.725
Valid N	67				

Based on Table 2, the mean and standard deviation of OPT scores were 47 and 1.725, respectively. Based on the OPT results, those who scored from 46 to 55 were selected and administered as an intermediate level of proficiency learners. Accordingly, 45 out of 67 Iranian EFL learners were selected and divided into three equal (*n*=15) groups, as two experimental and one control group.

The first research question investigated the effect of the ChatGPT on the writing performance of Iranian EFL students. To find out the difference among the three groups of the study, their performance in the pretest and posttest was compared via One-way ANOVA. The results are shown in the following Table.

Table 3. The ANOVA results of writing pretest

Source of Variance	SS	df	MS	F	Sig.
Between Groups	515.43	2	260.72		
Within Groups	203.66	43	4.83	8.402	.003
Total	715.09	45			

Based on the results of ANOVA, since p>0.05, (F=58.19), there is no significant difference among the three groups concerning their writing performance in the pretest. Then, the groups were compared after the instruction in the writing posttest. Table 4 illustrates the results.

Table 4. ANOVA results of writing posttest

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	106.133	2	132.067	8.402	.003
Within Groups	520.850	43	90.962		
Total	606.983	45			

As indicated in Table 4, the results show that F (2,45) = 8.402, and p < .05 which signifies a significant difference in writing posttest scores among the three groups. Nonetheless, a primary limitation of ANOVA is that it does not offer insights into where the variance originates or its specific location (Cresswell, 2014). In fact, a post-hoc test was followed to make multiple comparisons between the three groups to identify which of them appeared more effective. Therefore, in order to determine the location of the difference, a post-hoc Tukey HSD test was conducted. Table 5 illustrates the results.

Table 5. Tukey HSD of three groups

(I)	(J)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
1	2	-2.900*	.947	.009	-5.18	-.62
	3	-3.000*	.947	.007	-5.28	-.72
2	1	2.900*	.947	.009	.62	5.18
	3	-.100	.947	.994	-2.38	2.18
3	1	3.000*	.947	.007	.72	5.28
	2	.100	.947	.994	-2.18	2.38

*. The mean difference is significant at the 0.05 level.
(1)= First Experimental Group
(2)= Second Experimental Group
(3)= Control Group

As presented in Table 5, the difference was found between the control group and the first experimental group since p=.009. Moreover, the results showed the difference between the first experimental group and the second experimental group due to the fact that p=.007. Accordingly, it is argued that the second experimental group outperformed the first experimental and the control groups, revealing the efficiency of the instruction via ChatGPT along with teacher input.

The second research question explored the perception of EFL students of using ChatGPT to correct their errors. The data was analyzed using thematic analysis. Boyatzis (1998) defines thematic analysis as a method utilized in qualitative research to identify patterns and themes within a specific context, which entails systematically examining various phenomena to reveal categories and themes. The codes were categorized into two primary themes: positive and negative. Table 6 illustrates the three themes that emerged from the data's thematic analysis.

Table 6. Results of the interview

Categories	Themes	Example
EFL Learners' Perceptions before Using ChatGPT	a. Usefulness and availability of ChatGPT	"ChatGPT is good for practicing my writing skill since I can use it anytime from my PC"
Positive Perceptions	a. Time and Place Saving b. Ease to Work with c. Administration in Educational system d. Enjoyment and a new experience e. Providing Feedback	"I think ChatGPT is useful to save time for correcting English language errors by providing appropriate feedback. Without the tool, I do not know which part of the sentence is ungrammatical so I become confused."
Negative Perceptions	a. Technical Problems b. Easy to Cheat c. inability to use ChatGPT	1."ChatGPT is effective and convenient for learning English, particularly for me, as it offers tips and tricks that have enhanced my English skills. However, a serious problem was logging in." 2. "Although this AI tool has simplified many tasks, it has also spread plageriasm. You can easily write any essay in a very short period of time using it". 3. "I think that I need a lot of time and help to be able to use this tool".

According to Table 6, EFL students believed that ChatGPT was available for most of the learners. Furthermore, almost all participants had a positive view toward integrating ChatGPT in language teaching. An interviewee reported that

I did not think that AI could help me so much in learning a second language. It is very interesting that with the help of this tool, I can improve my writing.

They also appeared to be pleased with how ChatGPT assisted them in generating ideas and provided targeted information on the topic, allowing them to work independently. As one EFL learner highlighted,

The best aspect of ChatGPT is its ability to provide the information you request. You can request details to any level you wish. This has minimized the amount of time we spend thinking to gather ideas and information.

They also perceived that ChatGPT is useful for practicing different skills, especially writing. This means that the ChatGPT gave learners the chance to practice language on their own otherwise they may not be able to do that under normal circumstances. This was echoed in one interviewee's report:

I thought that someone should help me to learn to write in English. But now I see that I can improve without getting help from others or teachers.

However, one serious problem with the ChatGPT activities was logging in. As one interviewee stated,

Since you need a VPN for ChatGPT you have a great problem. I have personally problem with all types of VPNs and that is very disappointing.

Another interviewee noted that ChatGPT might result in diminished motivation to engage in thinking and an increased reliance on machines. She reported that

Sometimes I think that instead of writing a text, I should ask ChatGPT to do it for me. It is really tempting. I often worry that relying too much on AI could prevent me from writing independently. What if I end up not wanting to write by myself?

Another concern that was surprisingly expressed by 9 interviewees was that they do not possess the essential skills to utilize ChatGPT effectively to assist them in writing. An interviewee commented that

I think that it is not possible for me to learn to work with it [ChatGPT] with just a few training sessions and I need the help of a teacher.

Another EFL student said that

I presumed that with some prompts and direction, I could utilize ChatGPT to create content related to the course but now I think that it needs time, patience, practice, and training.

5. Discussion

This study was set up to explore the effect of ChatGPT on the EFL learners' writing performance. In addition, it investigated the Iranian EFL learners' perceptions of ChatGPT and its effect on correcting errors. The findings revealed that the ChatGPT combined with teacher feedback had a significantly positive effect on the EFL learners' writing performance. The results also showed that Iranian EFL learners favored ChatGPT, especially for time saving, ease to work with, administration in the educational system, enjoyment, and providing feedback.

Although no research has investigated the impact of combined teacher and ChatGBT on EFL learners' writing skill, the results of the present study are in line with a number of previous studies on the efficacy of ChatGPT in language teaching. Polakova and Ivenz (2014) explored the effectiveness of ChatGPT feedback in promoting the writing skills of EFL students. As reported, there was a significant improvement in various aspects of the participants' writing. The current results may also mirror the findings reported by Al-Mansour and Al-Shorma (2012), which suggest that incorporating ChatGPT resources with the traditional teaching method positively impacts student achievement. Furthermore, this outcome aligns with the conclusion drawn by Al-Jarf (2022), who discovered that using ChatGPT as an adjunct in conventional classrooms significantly influences writing structure. The findings of the current study align with those reported by Zaki (2023), who explored the impact of ChatGPT on the performance of Jordanian EFL students in English. In her research, the students in the experimental groups completed tasks utilizing dynamic animated illustrations through ChatGPT. In contrast, the control group received training with printed materials. Her results indicated that significant statistical differences existed between the control and experimental groups. ChatGPT should be based on various individual cognitive styles.

The qualitative section of the study revealed that students have a positive outlook on utilizing ChatGPTs for educational purposes, perceiving them as beneficial tools for language acquisition. Nonetheless, they stressed that ChatGPT should be regarded as a supplementary resource rather than a primary educational instrument. There were also apprehensions regarding the creation of language learning-oriented ChatGPTs due to the intricacies involved in mimicking language learning processes. When outlining the attributes, they would prefer in ChatGPTs intended for educational use, students highlighted the importance of having both collaborative and competitive features, especially for those in higher education. They expect ChatGPT to be engaging and pertinent to their learning experiences. Most participants explicitly stated or suggested that the objectives of ChatGPTs utilized for learning ought to be closely aligned with the learning outcomes of their curriculum, as indicated in earlier research (Estriegana et al., 2019).

It should be noted here that most of the participants in the interview mentioned the quality of the ChatGPT design to be motivating for the students, and make them engaged in class ativities.

The interview results are also consistent with those of Burston (2021) when he noted that "students perceived that ChatGPT use would make learning English easier" (p. 33). More specifically, those authors found "the greatest number of students mentioned enjoyment and motivation as potential advantages to using ChatGPTs in English-language learning" (Burston, 2021, p. 34). The findings also fit well with the findings of Alkamel and Alwagieh (2024) that revealed a positive view of Yemeni EFL learners towards using ChatGPT.

The results also align with the findings of the previous studies conducted on the ChatGPT. For example, Dizon and Gayed (2021) utilized a questionnaire using open-ended questions and discovered that a key theme from participants concerning the use of ChatGPT for educational purposes was 'fun'. They identified that enjoyment was a significant factor influencing the repeated use of ChatGPT and engaging with it in learning. It's reasonable for them to view enjoyment, pleasure, and fun derived from ChatGPT as crucial elements in determining its adoption for educational use. One of the main advantages of ChatGPT was the opportunity it offered teachers to design, implement, and reflect on the execution of a lesson that incorporated technology within a real classroom setting.

Additionally, based on the results of the interview the EFL learners reported some major concerns regarding the use of ChatGPT for improving writing. As they noted, technical problems, the possibility of providing the chance to cheat, and students' inability to use ChatGPT were the major pitfalls. This aligns with various studies conducted in this regard (e.g., Rudolph et al., 2023; van Dis et al., 2023). As Barrot (2023) notes, ChatGPT enables students to finish a writing task with minimal effort. Furthermore, as echoed in the literature, students' dependence on ChatGPT may compromise some crucial writing pedagogical objectives, such as encouraging creativity, cultivating critical thinking, and improving students' ability to put down their ideas with precision (Yanning, 2017). Khoso et al. (2023) have also outlined the common worries that students have regarding the use of ChatGPT, which encompass distractions and excessive dependence, the reliability of information, ethical considerations and issues of plagiarism, as well as concerns about privacy and security.

6. Conclusion

The present study aimed to explore how using ChatGPT's feedback combined with teacher feedback as a tool that affects the writing skills of EFL students. This could be one of the earliest practical studies looking into the impact of combined teacher and ChatGPT feedback on EFL students' academic writing abilities. The study, conducted as a mixed methods intervention, yielded results that were in line with the expectations. The notably positive influence, combined with students' favorable perception, could offer new insight into the existing literature on ChatGPT and other Automated Writing Evaluation (AWE) tools. The study's findings contribute to the AWE literature, particularly regarding the application of generative AI. They support and advance the understanding of feedback as a dialogic tool and ChatGPT as a constructive feedback mechanism that can be integrated into EFL writing classes considering that students have access to portable devices and the Internet. However, it should be noted that while AI tools such as ChatGPT may present assistance and immediate feedback currently they may not fully substitute human teachers' expertise (Golonka et al., 2014); therefore, human supervision is regarded to be crucial for maintaining the accuracy and integrity of produced content (Huang et al., 2023).

An important fact that should be mentioned is that individuals, particularly young people in Iran, are highly interested in utilizing technology. Furthermore, one of the most important objectives of the Iranian Ministry of Education has been to equip schools with computer-based tools and to facilitate effective work for both teachers and students in technologically advanced learning environments (Dashtestani & Hojatpanah, 2020). Nevertheless, without appropriate planning and training for the implementation of ChatGPT, the integration of such technologies in EFL classrooms may not have the desired impact. In addition, EFL teachers' attitude and their ability to use the tool in the classroom is of utmost importance as Mahapatra (2024) put it.

As Polakova and Ivenz (2024) argue, while ChatGPT provides exceptional opportunities for customized learning, educators should be cautious about potential drawbacks, such as depending too much on automated feedback and the possibility of plagiarism. Similar concerns about the incorporation of AI are articulated in the literature (e.g., Dai & Liu, 2024), emphasizing the importance of using AI in a balanced manner and maintaining human oversight to uphold academic honesty. This alignment with existing research reinforces the notion that while AI tools like ChatGPT can be extremely advantageous, they should not entirely replace human feedback or the collaborative elements of traditional learning settings.

There are implications based on the results. The findings indicate that EFL writing instructors should consider integrating ChatGPT into their teaching methods. By integrating AI feedback cycles and turning writing tasks into games, teachers can promote students' motivation and engagement in writing. This method can help boost students' self-efficacy through immediate, useful feedback. Using ChatGPT in group projects and peer review sessions can also encourage collaborative writing, as students can refine their drafts with AI-generated feedback before discussing them with peers. In AI-assisted courses, teachers should complement ChatGPT's feedback with personalized learning environment. A student-centered approach, with personalized learning paths and customizable feedback, can more effectively address individual writing challenges. By implementing these approaches, teachers can utilize ChatGPT's potential to improve students' writing skills while promoting a fair and ethical learning environment.

Like any research, the present study has its limitations, having been carried out within a specific context and with a small participant pool, which constrains the generalizability of the results. The dependence on self-reported data via questionnaires and interviews may lead to bias, as participants' answers could be swayed by social desirability and other influences. Moreover, the questionnaire could be enhanced by adding more items to reflect a wider array of students' experiences and preferences. The interview sample was minimal, comprising only ten participants from the ChatGPT group. To develop a more thorough understanding, additional analysis should be performed to investigate the feedback students obtained from both ChatGPT and another writing course, addressing aspects like grammar, vocabulary, and sentence structure. Future research could also examine the impact of ChatGPT on EFL students' writing fluency, complexity, and accuracy, offering a deeper insight into its distinct advantages or drawbacks in this context.

Moreover, since the intervention took only eight sessions further studies can use a longer intervention period to examine ChatGPT's effects. Another interesting area could be the effect of ChatGPT on various writing genres.

7. References

Abdullayeva, M., & Musayeva, Z. M. (2023). The impact of chat Gpt on students' writing skills: an exploration of AI-assisted writing tools. *International Conference of Educational Research and Innovation. 4*, 61–66.

Ahmed, P.H. (2019). Major writing challenges experienced by EFL learners in Soran University. *Journal of University of Human Development,5*(3), 120-126. https://doi.org/10.21928/juhd.v5n3y2019.pp120-126

Alenizi, M. A. K., Mohamed, Amr M., Shaaban, T. S. (2023). Revolutionizing EFL special education: how ChatGPT is transforming the way teachers approach language learning. *Innoeduca. International Journal of Technology and Educational Innovation,9*(2), pp. 5-23.

Al-Jarf, R. S. (2004). *Effect of online learning on struggling ESL college writers.* In: NECC 2002: National Educational Computing Conference Proceedings (23rd, San Antonio, Texas, June 17-19, 2002).

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

Al-Mansour, N. S., & Al-Shorman, R. A. (2012). The Effect of Computer Assisted-Instruction on Saudi University Students' Learning of *English. Journal of King Saud University–Languages and Translation, 24*(1), 51-56. http://dx.doi.org/10.1016/j.jksult.2009.10.001

Ausat, A., Massang, B., Efendi, M., Nofirman, N., & Riady, Y. (2023). Can Chat GPT Replace the Role of the Teacher in the Classroom: A Fundamental Analysis. *Journal on Education*, *5*(4), 16100-16106. https://doi.org/10.31004/joe.v5i4.2745

Barrot, J. S. (2023). Using ChatGPT for second language writing: pitfalls and potentials. *Assessing Writing, 5(*7), 67-88:100745. https://doi.org/10.1016/j.asw.2023.100745

Beatty, K. (2003). *Teaching and researching computer-assisted language learning.* Longman.

Boyatzis, R. (1998). *Transforming qualitative information: Thematic analysis and code development*. Sage Publications.

Burston, A. R. (2021). Exploring the potential of using an AI language model for automated essay scoring. *Research Methods in Applied Linguistics, 2*(2), 44-59. https://doi.org/10.1016/j.rmal.2023.100050

Cao, S., Zhou, S., Luo, Y., Wang, T., Zhou, T., & Xu, Y. (2022). A review of the ESL/EFL learners' gains from online peer feedback on English writing. *Frontiers in Psychology, 13,* 1035803. https://doi.org/10.3389/fpsyg.2022.1035803

Chen, X., Zou, D., Xie, H., Cheng, G., & Liu, C. (2022). Two Decades of Artificial Intelligence in Education: Contributors, Collabo-rations, Research Topics, Challenges, and Future Directions. *Educational Technology & Society, 25*(1), 28–47.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Sage.

Dai, K., & Liu, Q. (2024). Leveraging artificial intelligence (AI) in English as a foreign language (EFL) classes: Challenges and opportunities in the spotlight, *Computers in Human Behavior, 159.* https://doi.org/10.1016/j.chb.2024.108354

Dashtestani, R., & Hojatpanah, S (2020): Digital literacy of EFL students in a junior high school in Iran: voices of teachers, students and Ministry Directors. *Computer Assisted Language Learning*, https://doi.org/10.1080/09588221.2020.1744664

Dizon, G., & Gayed, J. (2021). Examining the impact of Grammarly on the quality of mobile L2 writing. *The JALT CALL Journal, 17*(2), 74–92. https://doi.org/10.29140/jaltcall.v17n2.336

Dwivedi, Y., j Sharma, A., Rana, N.P., Giannakis, M., Goel, P., Dutot, V. (2023). Evolution of artificial intelligence research in Technological Forecasting and Social Change: Research topics, trends, and future directions. *Technological Forecasting and Social Change, 192*. https://doi.org/10.1016/j.techfore.2023.122579

Estriegana, R. Medina, A. J. & Barchino, R.P. (2019). Student acceptance of virtual laboratory and practical work: An extension of the technology acceptance model. *Computers & Education, 3*(1), 33-48. https://doi.org/10.1016/j.compedu.2019.02.010

Farahian, M., & Ebadi, S. (2023). Collaborative digital writing and metacognitive knowledge in writing among TEFL students: the mediating role of online knowledge sharing. *Research and Practice in Technology Enhanced Learning, 18*(5), https://doi.org/10.58459/rptel.2023.18005

Fitria, T. N. (2021). Grammarly as AI-powered English writing assistant: Students' alternative for writing English. *Metathesis: Journal of English Language, Literature, and Teaching, 5*(1). https://doi.org/10.31002/metathesis.v5i1.351

Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2014). Technologies for foreign language learning: A review of technology types and their effectiveness. *Computer Assisted Language Learning, 27*(1), 70-105. https://doi.org/10.1080/09588221.2012.700315

Graham, F. (2022). Daily briefing: Will ChatGPT kill the essay assignment? *Nature.* https://doi.org/10.1038/d41586-022-04437-2

Grammarly. (2020). *About Grammarly*. Retrieved from https://support.grammarly.com/hc/en- us/categories/115000018611-About- Grammarly

Hakiki, M., Fadli, R., Samala, A. D., Fricticarani, A., Dayurni, P., Rahmadani, K., Astiti, A. D., & Sabir, A. (2023). Exploring the impact of using Chat-GPT on student learning outcomes in technology learning: The comprehensive experiment. *Advances in Mobile Learning Educational Research*, 859–872. https://doi.org/10.25082/AMLER.2023.02.013

Han, T., & Sari, E. (2022). An investigation on the use of automated feedback in Turkish EFL students' writing classes. *Computer Assisted Language Learning.* Advance online publication. https://doi.org/10.1080/09588221.2022.2067179

Harvey-Scholes, C. (2018). Computer-assisted detection of 90% of EFL student errors. *Computer Assisted Language Learning, 31*(1), 1-13. https://doi.org/10.1080/09588221.2017.1392322

Hasanah, Y. A., & Fatimah, S. (2020). *Undergraduates' ability in writing paraphrase and their perceptions of plagiarism*. Conference: 7th International Conference on English Language and Teaching (ICOELT 2019). https://doi.org/10.2991/assehr.k.200306.019

Hayati, A.M. (2005). The Computer and Language Teaching. *The Journal of Asian EFL, 4*, 1-8.

Huang, F., Kwak, H., & An, J. (2023). *Is ChatGPT better than human annotators? Potential and limitations of ChatGPT in explaining implicit hate speech*. arXiv [Preprint]. https://doi.org/10.1145/3543873.3587368

Jacobsen, D. M., & Lock, J.V. (2005). Technology and teacher education for a knowledge era: Mentoring for student futures, not our past. *Journal of Technology and Teacher Education*, 12 (1), pp. 75-87.

Jagdishbhai, N., & Thakkar, K. Y. (2023). Exploring the capabilities and limitations of GPT and Chat GPT in natural language processing. *Journal of Management Research and Analysis, 10*, 18–20. https://doi.org/10.18231/j.jmra.2023.004

Jahanban-Isfahlan, H., Hadidi Tamjid, N., & Seifoori, Z. (2023). Educational technology in Iranian high schools: EFL teachers' attitudes, perceived competence, and actual use. *Education Research International*, https://doi.org/10.1155/2017/9738264

Jiang, L., & Yu, S. (2022). Appropriating automated feedback in L2 writing: Experiences of Chinese EFL student writers. *Computer Assisted Language Learning, 35*(7), 1329–1353. https://doi.org/10.1080/09588221.2020.1799824

Johnson, K. Pasquale, F. & Chapman, J. (2019). *Artificial intelligence, machine learning, and bias in finance: Toward responsible innovation*. Fordham Press. Available at: https://ir.lawnet.fordham.edu/flr/vol88/iss2/5

Katar, O., Özkan, D., GPT, Yildirim, Ö., & Acharya, U. R. (2023). Evaluation of GPT-3 AI language model in research paper writing. *Turkish Journal of Science and Technology*, 18, 311–318. https://doi.org/10.55525/tjst.1272369

Khazaie, S., Hayati, A. M., & Jalilifar, A. R. (2013). Contribution to mobile-enhanced English language pedagogy among Iranian L2 learners. *Journal of English Language Teaching and Learning, 12*, 63-93.

Khoso, F. J., Ali, N., & Aslam, N. (2023). Use of Chat-GPT and AI tools by undergraduates: Students and teachers' perspective. *Spry Contemporary Educational Practices, 2*(2), 215–238.

Li, Y. (2024). Intelligent english teaching based on the pedagogy of performing another culture and ChatGPT technology. *Frontiers in Education Technology, 7*, 13. https://doi.org/10.22158/fet.v7n1p13

Lutai, N., & Besarab, T. (2020). Teaching paraphrasing in second language classrooms. *Naukovì Zapiski Nacìonal'nogo Unìversitetu «Ostroz'ka Akademìâ». Serìâ «Fìlologìâ», 9*(77), 224–227. https://doi.org/10.25264/2519-2558-2020-9(77)-224-227

Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: a mixed methods intervention study. *Smart Learn. Environ. 11*, 9. https://doi.org/10.1186/s40561-024-00295-9

Matsuda, P. K. (1998). Situating ESL writing in a cross-disciplinary context. *Written Communication*, *15*, 99-122. https://doi.org/10.1177/0741088398015001004

Meskill, C. (1999). 20 minutes into the future. In J. Egbert & E. Hanson-Smith (Eds.), *CALL environments: Research, Practice, and Critical Issues* (pp. 459-469). TESOL, Inc

Moghadam, M.H. (2023). Artificial Intelligence and the future of university education in Iran. *Quarterly Journal of Research and Planning in Higher Education, 29*(1). https://doi.org/10.61838/irphe.29.1.1

Muller, E. & Gregoric, A. (2017). Impact of ChatGPT on learning motivation: Teachers and students' voices. *Journal of English Studies in Arabia Felix, 2(1), 41–49.* https://doi.org/10.56540/jesaf.v2i1.51

Nguyen, T.T.L., & Suwannabubpha, S. (2021). EFL writing at Thai secondary schools: Teachers and students' views, difficulties and expectations. *Language Related Research, 12*(3), 187-214. https://doi.org/10.29252/LRR.12.3.7

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching, 6*(1), 1 – 22. https://doi.org/10.37074/jalt.2023.6.1.9

Patty, J. (2014). The use of AI in language learning: what you need to know. *Jurnal Review Pendidikan Dan Pengajaran, 7*(1), 642–654. https://doi.org/10.31004/jrpp.v7i1.24609

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

Rahimi, M., & Yadollahi, S. (2011). Foreign language learning attitude as a predictor of attitudes towards computer-assisted language learning. *Procedia Computer Science, 3,* 167-174. http://dx.doi.org/10.1016/j.procs.2010.12.029

Ram, B. & Verma, P. (2023). Artificial intelligence AI-based Chatbot study of ChatGPT, Google AI Bard and Baidu AI. *World Journal of Advanced Engineering Technology and Sciences, 8*(01), 258–261. https://doi.org/10.30574/wjaets.2023.8.1.0045

Rovikasari, M., S., & Supriyadi, S. (2020). Teachers' attitudes toward discovery learning: A case study in writing class. International *Journal of Educational Research Review, 5(*2), 135-140. https://doi.org/10.24331/ijere.686679

Rudolph, J. Tan, S. Tan, S. (2022). ChatGPT: bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning Teaching, 6* (1), 1–22. https://doi.org/10.37074/jalt.2023.6.1.9

Shidiq, M. (2023). The use of artificial intelligence-based chat-GPT and its challenges for the world of education; from the viewpoint of the development of creative writing skills. *Proceeding International Conference on Education, Sococity, and Humanity, 1* (1), 353-357.

Slamet, J. (2024). Potential of ChatGPT as a digital language learning assistant: EFL teachers' and students' perceptions. *Discov Artif Intell* 4, 46. https://doi.org/10.1007/s44163-024-00143-2

Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing, 57*, 100752. https://doi.org/10.1016/j.asw.2023.100752

Tafazoli, D. (2024). Exploring the potential of generative AI in democratizing English language education. *Computers and Education: Artificial Intelligence, 7*. https://doi.org/10.1016/j.caeai.2024.100275

van Dis, E. A., Bollen, J., Zuidema, W., van Rooij, R., & Bockting, C. L. (2023). ChatGPT: five priorities for research. *Nature, 614*(7947), 224 – 226. https://doi.org/10.1038/d41586-023-00288-7

Wambsganss, T., Janson, A., & Leimeister, J. M. (2022). Enhancing argumentative writing with automated feedback and social comparison nudging. *Computers & Education, 191,* 104644. https://doi.org/10.1016/j.compedu.2022.104644

Wang, Y. H., & Liao, H. C. (2008). The application of learning portfolio assessment for students in the technological and vocational education system. *Asian EFL Journal*, *10*(2), 132-154.

Wang, B., Hu, S. J., Sun, L., & Freiheit, T. (2020). Intelligent welding system technologies: State-of-the-art review and perspectives. *Journal of Manufacturing Systems, 56*, 373–391. https://doi.org/10.1016/j.jmsy.2020.06.020

Wilkens, O. (2020). Artificial intelligence in the workplace – A double-edged sword. *International Journal of Information and Learning Technology*, https://doi.org/10.1108/IJILT-02-2020-0022

Xue, Y & Wang, Y. (2022). Artificial intelligence for education and teaching. *Wireless Communications and Mobile Computing, 10*(6),1-10 https://doi.org/10.1155/2022/4750018

Yanning, D. (2017). Teaching and assessing critical thinking in second language writing: An infusion approach. *Chinese Journal of Applied Linguistics, 40*(4), 431 – 451. https://doi.org/10.1515/cjal-2017-0025

Yusriadi, Y., Rusnaedi, Siregar, N. A., Megawati, S., & Sakkir, G. (2023). Implementation of artificial intelligence in Indonesia. *International Journal of Data and Network Science, 7*(1), 283– 294. https://doi.org/10.5267/j.ijdns.2022.10.005

Zaki, N., Turaev, S., Shuaib, K., Krishnan, A., & Mohamed, E. (2023). Automating the mapping of course learning outcomes to program learning outcomes using natural language processing for accurate educational program evaluation. *Education and Information Technologies, 28*(12), 16723–16742. https://doi.org/10.1007/s10639-023-11877-4

Zolfaghari , F. (2023). Answer changing in online and traditional pen-paper tests: The case of upper intermediate EFL learners with different cognitive styles. *Applied Research on English Language*, *12*(3), 1-22. https://doi.org/10.22108/are.2023.138092.2112

8. Appendix A

Writing Pretest and Posttest

(Allotted Time: 30 Minutes)

Dear Participant,

Please write 150-200 words for each topic.

a. **My Family**

………
………
………
………
………………………………………………………………………………

b. **Historical Places of Iran**

………
………
………
………
………
……

c. **Your Favorite vacation**

………
………
………
………
……………………………………………………………

9. Appendix B

Semi-structured Interview

1. You have used ChatGPT to revise your texts. Tell us your experience with that.
2. How did you find working with ChatGPT? How did you feel?
3. How useful did you find ChatGPT? How did it help you?
4. Did you have any problems using ChatGPT? What are they?