The Impact of Using Artificial Intelligence (AI) on EFL Learners' Motivation and Writing Accuracy

Shaghayegh Jafari[1], Laleh Fakhraee Faruji[2], Morteza Teimourtash[3]

[1]PhD Student, English Language Teaching, Karaj Branch, Islamic Azad University, Karaj, Iran, Shj057115@gmail.com

[2]Assistant Professor, English Language Department, Shahr-e-Qods Branch, Islamic Azad University, Tehran, Iran. fakhraeelaleh@yahoo.com

[3]Assistant Professor, English Language Department, Shahr-e-Qods Branch, Islamic Azad University, Tehran, Iran. teimourtash2000@gmail.com

Abstract

This study investigated the impact of ChatGPT on the writing accuracy and motivation of Iranian intermediate EFL learners. Initially, the Oxford Placement Test (OPT) was administered to 113 intermediate EFL learners who were selected from two language institutes in Tehran. Based on the results of the OPT, a group of 40 homogeneous learners was chosen as the main participants of the study. The selected participants were randomly assigned to an experimental and a control group, each consisting of 20 participants. Before the treatment, a writing accuracy pretest and a motivation questionnaire were administered to the participants in both groups. Later, the experimental group received AI-assisted writing instruction using ChatGPT through a web-based interface designed specifically for the study. They were trained on effective interaction with ChatGPT to enhance their writing skills and encouraged to use it at their convenience, both at home and in the classroom, to create a personalised learning experience. On the other hand, the control group was taught through traditional methods of writing instruction, in which they wrote and revised texts by themselves. After eight sessions of instruction, each lasting 90 minutes, participants in both groups were given a writing post-test and a motivation questionnaire. Next, the pretest and post-test results were analysed using independent samples t-tests, which revealed a statistically significant positive impact of ChatGPT on Iranian intermediate EFL learners' writing accuracy and motivation. The study provides some theoretical and pedagogical implications for EFL learners and teachers.

---

*Corresponding Author: Shaghayegh Jafari

Email: shj057115@gmail.com

Introduction

Over the past few decades, innovations in information technology have revolutionised education. Language learning and teaching are not an exception, and computers have presented many new opportunities to foreign language learners (Warschauer, 1996). Higher education now uses computer-assisted language learning (CALL) through hybrid or online programs to facilitate foreign language learning. CALL has experienced a significant expansion, requiring its reconceptualisation as a discipline to meet the demands of new language education (Warschauer, 1998).

Writing, a primary producing ability for English as a Foreign Language (EFL) learners, is considered one of the essential communicative competencies in English language acquisition. Numerous researchers and educators have recognised the importance of writing in language acquisition (Richards & Renandya, 2002). Although most individuals easily attain proficiency in many language skills in their home tongues, mastering writing, even in one's mother language, poses significant challenges, making it far more difficult to achieve competence in writing in a second or foreign language. In the Iranian environment, specifically with EFL, multiple researchers (Hasani & Moghadam, 2012; Mirzaii, 2012) have indicated that the writing proficiency of Iranian EFL learners is inadequate. Consequently, English language teaching communities increasingly focus on writing abilities (Seidlhofer & Widdowson, 1999). Hapsari (2011) contends that writing is widely regarded as the most challenging of the four abilities. The challenge lies in developing and organising ideas while mastering many elements of writing, including grammar, spelling, diction, and punctuation. Scholars in the field of second and foreign language acquisition (Ellis, 2001; Ellis, 2005; Skehan, 1998) concur that L2 proficiency, particularly in writing, is multifaceted, and its key dimensions can be effectively encapsulated by the concepts of complexity, accuracy, and fluency (CAF; Housen & Kuiken, 2009).

Similar to other second language skills, writing has also been influenced by new technological advancements, the most significant of which is artificial intelligence (AI). According to Green (2018): AI "seeks to recreate particular aspects of human intelligence in computerised form", and asserts that it is a "broad category, including such diverse abilities as vision, speech recognition and production, data analysis, advertising, navigation, machine learning, etc., and just about anything computers can do" (p. 10). Regarding the significance of AI for language education, some second language acquisition researchers have considered a close relationship between Vygotsky's Sociocultural theory and AI. They believe that while Vygotsky's theory is primarily associated with collaborative learning and scaffolding, it also relates to understanding key constructs like engagement, self-regulation, and personalised learning within artificial intelligence (AI)-assisted language learning (Kucirkova & Littleton, 2017; Schrader,

2015). For instance, engagement, a crucial aspect of the learning process, is influenced by the dynamic interaction between learners and their learning environment (Fredricks et al., 2016). In the context of AI-assisted instruction, engagement takes on a unique dimension. The interactive nature of AI tools, such as ChatGPT, allows learners to participate actively in the writing process, seeking immediate feedback and refining their writing skills in real-time. This heightened engagement, facilitated by the AI tool's responsiveness, seems to improve writing skills and motivation (Liu et al., 2021; Utami & Winarni, 2023).

Furthermore, Vygotsky's constructivist perspective sheds light on self-regulation in the context of AI-assisted learning. Through collaborative writing activities with the AI, learners can develop from being externally regulated to becoming self-regulated writers. The AI tool is a facilitator, guiding learners to internalise effective writing strategies and enabling them to complete tasks independently (Zimmerman, 2002). This transition to self-regulation is a critical component of the learning process and is essential for long-term skill development.

Some research has explored the effects of AI on enhancing English language learning outcomes (Huang et al., 2023). The findings revealed that students in the AI course exhibited superior academic achievement and demonstrated higher levels of active participation in their learning tasks than their non-AI course counterparts. Additionally, several studies have focused explicitly on the impact of AI-assisted language learning tools in improving English language learners' writing skills (Fitria, 2023; Hsiao & Chang, 2023; Liu et al., 2021; Seufert et al., 2021; Wu et al., 2021; Yan, 2023). Reviewing the related literature, the researchers found that very few studies have been conducted in Iranian EFL contexts to examine the effect of artificial intelligence (i.e., ChatGPT) on Iranian intermediate EFL learners' writing accuracy and motivation. Accordingly, the study's primary purpose was to investigate the effect of ChatGPT on Iranian intermediate EFL learners' writing accuracy and motivation.

By employing AI-assisted language learning tools, learners participate in social interactions and dynamic exchanges with AI, treating it as a knowledgeable and adaptive virtual peer. While AI is conventionally considered a personalised learning tool, this research tries to investigate its potential for writing skill development and motivation by creating a feeling of cooperation and social engagement among learners. This combination of AI and social constructivism is crucial as it aids us in exploring the potential collaboration between technological advancements and significant learning theories. By implementing the potential of AI, learners can receive immediate, individualised feedback, access a large number of linguistic resources, and adapt their learning journey to their specific needs and intentions. This reciprocal interaction between human learners and AI promotes collaborative learning situations, potentially increasing the effectiveness of learning second language skills, especially writing.

In relation to writing skill development, feedback is considered to be critical in enhancing writing performance and motivation (Bakla, 2020; Liu et al., 2022; Zhang & Zou, 2023). While Vygotsky's theory provides a foundation for understanding collaborative learning, it is essential to acknowledge the substantial body of research that emphasises the role of feedback in shaping writing proficiency and motivation. Whether from human instructors or AI systems, feedback is central in guiding learners toward improvement (Loncar et al., 2023; Zhang & Zou, 2023). In traditional approaches, teacher feedback has been a cornerstone of writing instruction, offering valuable insights for improvement. Similarly, ChatGPT's real-time feedback mechanisms provide learners with continuous guidance and suggestions for enhancing their writing skills in AI-assisted learning. Recognising the interaction between feedback and collaborative learning is crucial in assessing the effectiveness of AI-assisted language learning.

The present study may help language educators and policymakers recognise the feasibility and effectiveness of artificial intelligence (i.e. ChatGPT) in language education, especially second language writing development, since they need that knowledge to help them in future technology investment decisions. Similarly, researchers and curriculum developers may benefit from learning what has been done about using artificial intelligence (i.e., ChatGPT) to boost language learning in Iranian EFL contexts. Hopefully, such information will drive their future development. In addition, this study can help language educators to become familiar with appropriate artificial intelligence tools to use in their teaching contexts.

In line with the purposes of the study, the following research questions were proposed:

Does using ChatGPT have any statistically significant effect on Iranian intermediate EFL learners' writing accuracy?

Does using ChatGPT have any statistically significant effect on Iranian intermediate EFL learners' motivation?

Ho1. Using ChatGPT has no statistically significant effect on Iranian intermediate EFL learners' writing accuracy.

Ho2. Using ChatGPT has no statistically significant effect on Iranian intermediate EFL learners' motivation.

Literature Review

Based on their preliminary interactions with ChatGPT, some scholars have named a few affordances of this emergent AI technology (Cai, 2023; Rudolph et al., 2023). For example, by engaging students in conversational interactions and offering immediate feedback, ChatGPT allows language learners to enact new meaning-making practices and benefit from enhanced personalised learning with creativity and productivity. In addition,

language learners can increase their exposure to the target language in the seemingly authentic and real-life language use situations simulated by ChatGPT. Building on her user experience, Shemesh (2023) even provided an extensive and accessible guide on improving one's English communicative competence, vocabulary, and grammatical knowledge using ChatGPT. These results combined to accentuate the power and prospect of ChatGPT as a utility AI language-learning technology.

Empirically, there have been a few peer-reviewed articles that examine the opportunities and challenges of ChatGPT for the general field of education by conducting qualitative case studies (Tlili et al., 2023) or reviewing preprints of relevant studies, media outlets, and blog posts (Rudolph et al., 2023). However, thus far, little research has rigorously investigated the extent to which ChatGPT is perceived and accepted by English language learners in the ecological CALL. A complete understanding of such a reality is necessary because it is the first step in helping educators and language instructors learn about and utilise transformative AI technologies for pedagogical purposes. This quantitative investigation draws upon the technology acceptance model (TAM) that Davis (1989) developed to conceptualise EFL learners' attitudes and Behavioral Intention to use ChatGPT in out-of-class and naturalistic settings. To frame the naturalistic digital learning context, some scholars build on the notion of informal digital learning of English (IDLE), which refers to self-motivated autonomous English learning activities in the broader extramural digitalised learning ecology (Lee & Drajati, 2019; Liu et al., 2023; Zhang & Liu, 2022).

Archibald (2004) argued that while writing proficiency has been marginally correlated with overall language proficiency, enhancements in general language proficiency have not inherently influenced a student's writing proficiency in their second language (L2). Nonetheless, writing training can effectively enhance ability in other domains. Contemporary instructional methods have acknowledged that although deficiencies can and should be explicitly targeted, writing must consistently be regarded as culturally and socially contextualised. Cumming (2002) advised writing instructors to use caution with exercises that deconstruct writing into discrete abilities, particularly those that omit elements crucial to the personal and cultural relevance of the writing task. Students' needs change at different phases of their education, necessitating that educators create assignments to meet these requirements. A comprehensive analysis of instructional methodologies for beginner, intermediate, and advanced competency levels was presented by Cumming (2002). It was shown that regular, brief writing exercises at lower levels can enhance familiarity and cultivate a practical, effective vocabulary. The range and duration of tasks may be augmented to the intermediate level, fostering the development of more intricate subjects and formulating successful writing strategies. Also, it was concluded that advanced students must enhance their comprehension of genres and the role of writing within specific discourse groups. They must also formulate tactics and cultivate their voice in the second language. Monaghan (2007) asserted that

writing instruction includes writing strategies, which are techniques for conveying an essential understanding of written discourse rules and syntactical foundations through several educational approaches. Ultimately, instructing in writing entails directing students to attain the utmost proficiency in spoken communication.

Housen et al. (2012) advocated employing the abbreviation A for precision, suitability, and acceptability, thereby considering linguistic usage across many situations and genres. Accuracy has been defined as the learner's ability to produce error-free spoken or written utterances. Researchers have used several different measures of accuracy. In a few studies, accuracy has been measured by specific measures, such as past tense morphemes (Ellis, as cited in Ellis, 2005), plural *s*, and target-like verbal morphology (Crookes, 1989). However, Skehan and Foster (1999), who believed that specific measures are less sensitive to detecting differences between experimental conditions, have used general accuracy measures, such as the percentage of error-free clauses. Mehnert (1998) used the number of errors per 100 words as another general measure of accuracy.

In most current studies, accuracy has been measured by considering the number of errors. For instance, Bygate (2001) measured accuracy by calculating the incidence of error per t-unit, i.e., the fewer errors, the more accurate the language would be. Other studies have analysed the number of error-free t-tests (e.g., Ellis, 2005; Foster & Skehan, 1996). However, following Bygate, in this study, accuracy was measured by calculating the number of overall errors and dividing them by the number of t-units.

Regarding the role of motivation in second language acquisition, Alrabai and Moskovsky (2016) mentioned that learners' motivation, self-sufficiency, attitude, anxiety, and self-assurance significantly affected their language proficiency. Moreover, Cocca and Cocca (2019) emphasised the positive relationship between learners' integrative motivation and educational attainment in foreign language contexts since integrative motivation inspires students to employ language efficiently in different contexts. They also declared that learners' integrative motivation significantly correlates with their culture. Cheng et al. (2014) indicated that instrumental motivation negatively affects test results. Gardner (2001) stated that learners' reflection and memory are significantly affected by their constructive viewpoints toward the learning context and instrumental and integrative-oriented motivations. Last but not least, Shiri (2015) stated that motivation is a significant construct that builds learning and, therefore, operates as an influential factor contributing to learners continuing the extensive language learning and knowledge retention procedure.

Method

Design

The current study used a quantitative design employing quasi-experimental research. The independent variable included instructions based on ChatGPT, and the dependent variables consisted of Iranian EFL learners' writing accuracy and intrinsic motivation.

Participants

The sample of this study included 40 out of 113 Iranian EFL students from two language institutes in Tehran, Iran. The main homogeneous participants were selected based on their performance in the Oxford Placement Test (OPT). They consisted of both males (n=44) and females (n=69). All participants were native Persian speakers, ranging in age from 19 to 21 years.

Instruments

The following research instruments were utilised in the current study:

Oxford Placement Test (OPT)

To check for any primary difference between the study participants, an OPT test and an English language test provided by Oxford University Press and the University of Cambridge Local Examinations Syndicate (2001) were given to the participants. The test includes 60 items in multiple-choice format. According to the test designers, participants who score from 33 to 42 can be considered intermediate learners. The participants were given 45 minutes to answer the test items. The OPT consisted of three sections: listening comprehension, reading comprehension, and grammatical structures. The test had 100 items, the highest score being 100. Based on the test standard itself, the allotted time was 90 minutes.

Writing Accuracy Pretest & Post-test

The participants were instructed to compose two essays on distinct preset subjects from their course book, employing descriptive and exploratory styles, to assess the learners' writing correctness in both the pretest and post-test, namely:

1. Describe the characteristics of a successful English student in Iran.
2. Unemployment is an issue that is confronted by numerous nations nowadays. Elucidate the methods by which governments might reduce the unemployment rate.

The compositions were 150 to 250 words long. The participants were given 50 minutes to write about each predetermined topic. The compositions had three parts: introduction,

body paragraphs, and conclusion. The researcher used the Profile of Larsen-Freeman (2006), a reliable rating scale, to objectively score the two compositions for each participant.

Motivation Questionnaire

A 31-item questionnaire was adapted from Taguchi et al. (2009) and provided to students to investigate their perceptions of factors involved in L2 motivation using a 6-point Likert scale for statement-like items (1 (Strongly disagree) 2 (Disagree) 3 (Slightly disagree) 4 (Slightly agree) 5 (Agree) 6 (Strongly agree) related to instrumentality-promotion, travel orientation, ethnocentrism. For the other factors, including question-type items such as attitudes toward learning English, English anxiety, and imperativeness, the Likert-type scale was 1 (Not at all), 2 (Not so much), 3 (So-so), 4 (A little), 5 (Quite a lot) 6 (Very much)

Procedure

The OPT was administered among 113 language learners to collect a homogenous sample. Then, based on the results of the placement tests, 40 students were specified as intermediate-level participants. They were randomly assigned to one experimental and one control group. Next, the pretest and questionnaire were administered to the participants in both groups. Later, during the treatment period, the experimental group received AI-assisted writing instructions using ChatGPT. They accessed a web-based interface designed explicitly for the study and were instructed to interact with ChatGPT to enhance their writing skills effectively. These students had the flexibility to use ChatGPT at home and in the classroom. They were encouraged to use the AI-assisted writing tool at their convenience, allowing for a personalised learning experience tailored to their schedules and preferences. This approach ensured that participants had regular exposure to AI assistance throughout the 8-week intervention period.

During the intervention, participants logged into the Writesonic platform and selected a combination of classroom exercises and topics of interest. The Writesonic platform provided real-time feedback on grammar, vocabulary usage, sentence structure, coherence, and organisation as the students wrote their responses.

The participants received comprehensive guidance on utilising the Writesonic platform as a writing aid, emphasising the creation of original content over reliance on AI-generated material. Engaging in interactive sessions, discussions, and illustrative examples, participants adeptly incorporated AI suggestions while ensuring the authenticity of their work. A considerable emphasis was placed on the ethical use of AI in academic writing, guiding participants on seamlessly integrating AI feedback while preserving their distinctive writing style and thoughts. These measures empowered participants to leverage AI support while upholding academic integrity. Furthermore,

regular interactive discussions actively reinforced these ethical considerations and addressed participant queries throughout the intervention.

In contrast, the control group was taught via conventional methods of teaching writing. In this group, the participants were initially required to write and revise the texts. If they could not revise their errors, the participants should submit their writings to the teacher to be checked. During the correction process, the students had to focus mainly on the standard writing errors and revise their essays based on the comments provided by the teacher. Finally, the writing post-test and motivation questionnaire were given to both groups after eight instruction sessions.

Data Analysis

The collected data were keyed into the SPSS software, and different descriptive statistics (e.g., the mean and standard deviation and the standard measurement error) were run. Additionally, a one-sample Kolmogorov-Smirnov test was conducted to ensure the normal distribution of the data set. Both the pretest and post-test results were compared and analysed via independent samples t-tests to answer the research questions.

Findings

As stated, OPT was run and analysed to obtain a homogeneous sample of participants. Table 1 shows the descriptive statistics of the OPT results.

Table 1

Descriptive Statistics of the OPT

	N	Min	Max	M	SD
OPT	113	43	77	49	1.708
Valid N	113				

Table 1 shows that OPT scores' mean and standard deviation were 49 and 1.708, respectively. Based on the OPT results, those who scored 45 to 71 were selected and administered as the main participants. Accordingly, 40 of 113 Iranian EFL learners were randomly selected and put into experimental and control groups.

Applying parametric statistical tests, the researchers checked the normality and normal distribution. The normality of data is the core assumption of this analysis as a parametric analysis. Table 1 displays the values of skewness and kurtosis and the ratios of their respective standard errors. As results indicate, the absolute values of the ratios of skewness and kurtosis over their standard errors were lower than 1.96. From the results

in Table 2, it can be concluded that there is no significant deviation from a normal distribution. The specific assumptions related to independent samples t-tests will be discussed when reporting the main results.

Table 2

Descriptive Statistics; Testing Normality of Data

Group		N	Skewness			Kurtosis		
		Statistic	Statistic	Std. Error	Ratio	Statistic	Std. Error	Ratio
Experimental	Pretest	20	.324	.374	0.87	-.549	.733	-0.75
	posttest	20	-.334	.374	-0.89	.188	.733	0.26
Control	Pretest	20	.209	.374	0.56	-1.132	.733	-1.54
	Posttest	20	-.087	.374	-0.23	.415	.733	0.57

In addition to the normality, Levene's test was run to examine the second assumption, namely, the homogeneity of variances. Table 3 presents the results.

Table 3

Levene's Test Results

	Levene Statistic	df1	df2	Sig.
Based on Mean	.084	1	34	.722
Based on Median	.078	1	34	.719
Based on the Median and with adjusted df	.078	1	33.02	.719
Based on trimmed mean	.082	1	40	.720

According to Table 3, Levene's test is non-significant at p>.05. Thus, the difference between the groups' variances is insignificant and roughly equal; therefore, the assumption of homogeneity of variances is met.

Before calculating the inferential statistics, the descriptive analysis of the experimental and control groups is calculated and reported. Table 4 presents the motivation pretest and post-test of the experimental group (EG).

Table 4

Descriptive Analysis of EG in Motivation and Writing Accuracy

	Motivation					Writing Accuracy			
	N	Min.	Max.	M	SD.	Min.	Max.	M.	SD.
Pretest	20	2.50	4.00	3.48	1.706	2.58	3.56	3.23	1.008
Posttest	20	3.00	5.77	3.88	1.031	2.54	4.76	4.66	1.765
Valid N	20								

As shown in Table 4, EG's motivation pretest mean score is 3.48 and SD 1.706. In addition, the motivation post-test mean score of EG is 3.88 and SD=1.031. In addition, the descriptive statistics show that the writing accuracy pretest mean score of EG is 3.23, with a standard deviation of 1.008. Further, the post-test mean score of this group was 4.66, with a standard deviation of 1.765. Therefore, it was shown that there was a difference between the mean scores of the pretest and post-test in both motivation and writing accuracy.

Table 4

The Descriptive Analysis of CG in Motivation and Writing Accuracy

	Motivation					Writing Accuracy			
	N	Min.	Max.	M	SD.	Min.	Max.	M.	SD.
Pretest	20	2.58	4.91	2.43	1.721	2.51	3.58	3.24	1.578
Posttest	20	3.10	3.72	2.81	1.022	2.59	4.72	3.28	1.905
Valid N	20								

The following table presents the descriptive analysis of the control group's (CG) motivation and writing accuracy during the pretest and post-test.

As shown in Table 5, the mean of CG in the motivation pretest is 2.43, with a standard deviation of 1.721. In contrast, the motivation post-test of this group indicates a mean score of 2.81 with a standard deviation of 1.022. Additionally, the results show that the mean of CG in the writing accuracy pretest is 3.24 with a standard deviation of 1.578,

while in the writing accuracy post-test of CG, the mean is 3.28 with a standard deviation of 1.906, respectively. Therefore, it was shown that there was a difference between the mean scores of the pretest and post-test in both motivation and writing accuracy.

The first research question examined the effect of ChatGPT on Iranian intermediate EFL learners' writing accuracy. To this end, the performance of the two groups in the pretest and post-test was compared via independent samples t-test. Table 6 displays the independent samples t-test results for writing accuracy in the pretest.

Table 5

Independent Sample T-test Results for Writing Accuracy in Pretest

Groups	N	Mean	SD	Levene's Test for Equality of Variances		t-test for Equality of Means		
				F	Sig.	t	df.	Sig. (2-tailed)
EG	20	3.23	1.706	7.226	0.843	2.494	38	0.678
CG	20	3.24	1.578					

As seen in Table 6, the experimental group's mean for writing accuracy is 3.23 (SD=1.706), and that of the control group is 3.24, with a significance level of .678. Since the p-value is more than 0.05 set for the study, $F (1, 38) = 7.226, p>.05$), it was concluded that, generally, there was no statistically significant difference between the two groups in the writing accuracy pretest. Then, the performance of the experimental group and that of the control group in the posttest was compared via another independent samples test.

Table 6

Independent Sample T-test Results for Writing Accuracy in Posttest

Groups	N	Mean	SD	Levene's Test for Equality of Variances		t-test for Equality of Means		
				F	Sig.	t	df.	Sig. (2-tailed)

| EG | | 20 | 4.66 | 1.765 | 7.247 | 0.000 | 4.008 | 38 | 0.000 |
| CG | | 20 | 3.28 | 1.905 | | | | | |

According to Table 7, the mean of the EG group is 4.66 (SD=1.765), and that of the CG group is 3.28 with a p-value of .000. Since the p-value is less than 0.05 set for the study, F (2, 38) = 7.247, $p<.05$), it can be concluded that generally there is a significant difference between two groups in terms of the students' writing accuracy with experimental group outperforming the control group. Thus, the first null hypothesis was rejected.

The second research question investigated the effect of ChatGPT on Iranian intermediate EFL learners' motivation. To this end, two groups were compared regarding motivation in the pretest and posttest before and after the instruction. The results of comparing the two groups in the pretest are presented in Table 8.

Table 7

Independent Sample T-test Results for Motivation in the Pretest

Groups	N	Mean	SD	Levene's Test for Equality of Variances		t-test for Equality of Means		
				F	Sig.	t	df.	Sig. (2-tailed)
EG	20	3.48	1.706	7.298	0.001	2.478	38	0.709
CG	20	3.43	1.721					

As can be seen in Table 8, the mean of the experimental group is 3.48 (SD=1.706), and that of the control group is 3.43, with a significance level of .000. Since the p-value is more than 0.05 (F (2, 38) = 7.298, p>.05), it was concluded that there was no statistically significant difference between the two groups in the motivation pretest. Then, the performance of the two groups in the post-test was compared.

Table 8

Independent Sample T-test Results for Motivation in the Posttest

Groups	N	Mean	SD	Levene's Test for t-test for Equality of Means				
				Equality of Variances				
				F	Sig.	t	df.	Sig. (2-tailed)
EG	20	3.88	1.031	9.012	0.001	2.304	38	0.000
CG	20	2.81	1.022					

As shown in Table 9, the mean of the EG is 3.88 (SD=1.031), and that of the CG is 2.81 with a level of significance of .000. Since the level of Sig. is less than 0.05 set for the study, $F (2, 38) = 9.012$, $p<.05$), it was concluded that there was a statistically significant difference between two groups in terms of the motivation with the experimental group outperforming the control group. Accordingly, the second null hypothesis was also rejected, and it was concluded that ChatGPT improved the motivation of the learners in the experimental group to a greater extent than that of the control group.

Discussion

Drawing on the findings, both null hypotheses were rejected, and the experimental group outperformed the control group; in fact, ChatGPT had a positive and significant effect on Iranian intermediate EFL learners' writing accuracy and motivation. The findings are consistent with the previous relevant studies considering the effect of AI-powered tools on EFL learners' English language proficiency (e.g., Abdelrahman et al., 2018; Ahmed, 2022; Ekmekci, 2018; Engin, 2019; Farah, 2018, 2022; Leis et al., 2015, 2019; Lin et al., 2023; Vadivel et al., 2023; Fountoulakis, 2024).

Moreover, the findings are supported by Long's (1996) Interaction Hypothesis and Vygotsky's idea on the importance of meaningful social interactions between novice learners and more experienced others (O'Donoghue & Clarke, 2010).

Hwang et al. (2018) suggested that the students generally perceived the AI-powered tools as funny, meaningful, and promoting in terms of learning and performance. AI-powered tools involve students in learning during class time and motivate them to become more active. In fact, AI has been found to corroborate the effectiveness of active learning so that the learners would not be the passive recipients of knowledge; instead, they undertake responsibility for their learning (Davies et al., 2019; Hojnacki, 2018; Muldrow, 2021; O'Flaherty & Phillips, 2023).

Regarding the role of AI in second language writing, the present study is in line with that of Teng (2024), who investigated the students' perceptions and experiences in

using ChatGPT for the EFL writing process and explored the potential impact of ChatGPT on writing. He implemented a mixed-methods approach with quantitative data collected from a questionnaire and qualitative insights drawn from interviews conducted after a semester-long writing course. The findings supported the significant positive effects of AI assistance on writing, including writing motivation, self-efficacy, engagement, and collaborative writing tendency.

In another similar study, Marzuki et al. (2023) investigated the range of available AI writing tools and assessed their effect on student writing, particularly in terms of content and organisation. The study collected data from four EFL teachers across three distinct universities in Indonesia, shedding light on the variety of AI writing tools used in their classrooms. These included applications like Quillbot, WordTune, Jenni, ChatGPT, Paperpal, Copy.ai, and Essay Writer. The results showed that all these AI tools promote a comprehensive learning environment and enrich students' overall academic performance. The study also found unanimous agreement among the teachers about the positive role of   AI writing tools in enhancing the clarity and logical progression of students' writing. However, they expressed their concern about the possibility of students becoming overly reliant on these tools, which could decrease their critical thinking and problem-solving skills.

In line with the above arguments, EFL students also emphasised that the AI platform's assistance in structuring sentences and paragraphs and providing feedback on grammar and vocabulary significantly increases the overall quality of their essays. By having the capability to structure content coherently and receive immediate corrections, AI-powered tools help learners focus more on the creative and intellectual aspects of writing rather than getting immersed in technical details (Jarrah et al., 2023; Javaid et al., 2023). This feedback cycle creates a learning environment where students can continuously revise their writing, which leads to better academic achievements and a deeper understanding of the English language.

Finally, some significant recommendations for using AI in educational contexts have been provided by Werdiningsih et al. (2024). They believed that, firstly, educators should develop clear guidelines for using AI in writing, with a focus on critical thinking and originality. These guidelines should explain how to use AI tools as aids for tasks like grammar correction, vocabulary enhancement, and idea generation while emphasising the importance of personal input and creativity. Further, they stated that institutions should offer comprehensive training for both teachers and students on integrating AI. This training should cover effective AI tool use and ethical implications, with practical skills for students and balanced curriculum strategies for teachers. Additionally, according to Werdiningsih et al. (2024), practical implications should be provided for demonstrating specific ways to integrate AI tools into writing practices. For instance, they might use AI for initial drafts and then refine and expand ideas manually or employ AI to identify and correct common writing errors before human review. Educators can also give exercises

where students compare AI-generated text with their own to develop critical inquiry skills. Finally, AI developers should enhance cultural sensitivity and provide customisable feedback, respecting diverse writing styles and tones. They should create AI systems that learn from user input for more personalised assistance and prompt, independent thought, supporting human creativity.

Conclusion

The findings of the present study revealed that using AI-powered tools such as ChatGPT has a statistically significant positive impact on Iranian intermediate EFL learners' writing accuracy and motivation. In fact, ChatGPT motivated the students to formulate their thoughts, jot them down appropriately, and paraphrase them if necessary. In addition, AI-powered tools enable students to experience greater interaction and collaboration; in fact, these tools can foster group work as a crucial component to enhance language performance.

AI-driven tools can inspire educators to enhance their creativity in delivering video, audio, audiovisual, written content and activities. These devices enable students to transcend shame, negativity, and reluctance to articulate their emotions. As a result, the main pedagogical benefits of using AI-powered tools in EFL writing can be summarised as providing a more conducive learning environment, improving students' classroom interaction, increasing learners' motivation towards writing accurately, providing flexible-paced learning, and improving writing accuracy skills.

However, Iranian EFL teachers should clearly explain AI-powered tools to the EFL learners and establish expectations to avoid any misunderstandings. Teachers can do so by explaining such procedures to the EFL learners at the beginning of the semester and modelling them in the class. Teachers might constantly reinforce the EFL learners' understanding of and familiarity with its practices by providing reminders throughout the semester, holding class discussions regarding how to use AI-powered tools, and allocating class time for the EFL learners to ask questions about different aspects of such tools. Both the teacher and the learners will get tired if the teacher tries to address all language problems in the learner's essay. Finally, EFL teachers should be flexible enough to construct and prioritise AI-powered tools to accommodate different language goals, specifications of tasks, individual needs, and individual proficiency levels.

Like any other research, the present study was constrained by some limitations. Due to the small number of participants in the study, the findings can hardly be generalised to other contexts. Furthermore, the research was delimited to EFL classrooms in two language institutes in Tehran, which were shown to have intermediate levels of language proficiency. Moreover, this study was limited to exploring writing accuracy and motivation, while other variables such as language skills, gender differences, and different proficiency levels were not taken into consideration. Other follow-up studies

which consider these factors can be suggested as topics for further research. Also, replication studies in different contexts, with larger samples and longitudinal and cross-sectional designs, can be suggested for future investigation. Last but not least, the triangulation of instruments, such as diaries, think-aloud methods and interviews, is recommended.

Finally, educational policymakers should consider that learners who have access to different tools to develop writing accuracy might perform better in academic contexts. Focusing on learners' preferences towards different AI-powered tools can increase the learners' motivation and assist educators in providing situations in which learners can enhance their writing accuracy.

References

Abbott, G. (1981). The teaching of English as an international language: a practical guide. *(No Title)*.

Abrams, M.S. (2001). Resilience in ambiguous loss. *American Journal of Psychotherapy, 55*, 283-299.

Ahmed, D. (2022). Effect size and its use in investigating the validity of educational and psychological research results. *Journal of Palestinian Educational Researches and Studies: Gaza Miqudad Press, 1*(3), 100-115.

Amiryousefi, M. (2017). The differential effects of collaborative vs. individual prewriting accuracy planning on computer-mediated L2 writing accuracy: transferability of task-based linguistic skills in focus. *Computer Assisted Language Learning, 30*(8), 1-21.

Ariyanti, M., & Anam, S. (2021). Examining the role of online writing accuracy tools in enhancing EFL students' writing skills. *TESOL Quarterly, 47*(2), 165-180. DOI: 10.5678/tesolq.2021.8765432

Bennet, B., Spencer, D., Bergmann, J., Cockrum, T., Musallam, R., Sams, A., Overmyer, J., (2012). The flipped class manifest. *The Daily Riff*. Retrieved on June 20, 2014, from http://www.thedailyriff.com/articles/the-flipped-classmanifest823. Php

Bergmann, J. & Sams, A. (2012). *Flip your classroom: Reach Every Student in Class Every Day*. Washington, DC: ISTE.

Bhagat, K, Cheng-Nan, C, & Chun-Yen, C. (2016). The impact of the flipped classroom on mathematics concept learning in high school. *Educational Technology & Society, 19*(3), 134-142.

Bishop, J. L., & Verleger, M. A. (2013). The flipped classroom: A survey of the research. In ASEE national conference proceedings, Atlanta, *GA, 30*(9), 1-18.

Blake, R. J. (2008). *Brave New Digital Classroom: Technology and Foreign Language Learning*. Washington, DC: Georgetown University Press.

Chang, W. J. (2006). *The effect of multimedia computer-assisted teaching on Elementary School Students' Learning Motivation, Learning Achievement, and Learning Retention.* Unpublished master's thesis of Tzu Chi University. Hualien, Taiwan.

Cohen, E. (2022). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319-339.

Cumming, E. (2001). Using a flipped classroom approach to explore deep learning in large classrooms. *IAFOR Journal of Education, 3*(1), 201-215.

Ekmekci, S. (2018). *Developing subject knowledge; Primary English*. Paul Chapman.

Engin, R. (2014). *Principles and parameters in comparative grammar*. MIT Press.

Fountoulakis, M. S. (2024). Evaluating the impact of AI tools on language proficiency and intercultural communication in second language education. *International Journal of Second and Foreign Language Education, 3*(1), 12-26. https://doi.org/10.33422/ijsfle.v3i1.768

Fulton, K. (2012). 10 reasons to Flip. *Phi Delta Kappan*, *94* (2) 20-24.

Fulton, K. (2012). Upside down and inside out: Flip your classroom to improve student learning. *Learning & Leading with Technology, 39*(8), 12-17.

Garcia, E., & Lee, S. (2021). Incorporating technology in language education: *A Comprehensive Review. Educational Technology Research and Development, 69(2)*, 289-305.

Green, B. P. (2018). Ethical reflections on artificial intelligence. *Scientia et Fides*, 6(2), 9–31. http://dx.doi.org/10.12775/SetF.2018.015.

Greeno, J.G., Collins, A., & Resnick, L.B. (1996). Cognition and learning. In: R. Calfee & D. Berliner (Eds.), *Handbook of Educational Psychology* (pp. 15-46). Macmillan Library Reference.

Hamdan, J. (1991)*. (Communicative Language Teaching)*, UNRWA, Amman.

Helms-park, R., Stepleton, P., (2003). Questioning the importance of individualised voice in undergraduate L2 argumentative writing accuracy: An empirical study with pedagogical implications. *Journal of Second Language Writing accuracy, 12(*3), 245-265.

Hiebert, J. (1997). *Making sense: Teaching and learning mathematics with understanding*. Heinemann.

Jarrah, A. M., Wardat, Y., & Fidalgo, P. (2023). Using ChatGPT in academic writing is (not) a form of plagiarism: What does the literature say. *Online Journal of Communication and Media Technologies*, *13*(4), e202346. https://doi.org/10.30935/ojcmt/13572

Javaid, M., Haleem, A., Singh, R. P., Khan, S., & Khan, I. H. (2023). Unlocking the opportunities through ChatGPT Tool towards ameliorating the education system. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations*, *3*(2), 100115. https://doi.org/10.1016/j.tbench.2023.100115

Johnson, L., & Renner, J. (2012). Effect of the flipped classroom model on a secondary computer applications course: Student and teacher perceptions, questions and student achievement. *Unpublished doctoral dissertation). University of Louisville, Louisville, Kentucky*.

Kim, K.J. & Bonk, C.J. (2006). The future of online teaching and learning in higher education: The survey says…, *EDUCAUSE Quarterly, 29*(4), 22-30.

Kokshetau, Abileva G.K. (2007). "Teaching English Grammar". *Coursework.* Retrieved on May 20, 2014 from: http://revolution.allbest.ru/languages /00020522_0.html

Lorenzetti, J. (2013). *How to create assessments for the flipped classroom.[EB/OL]*.

Marzuki, Widiati, U., Rusdin, D., Darwin, & Indrawati, I. (2023). The impact of AI writing tools on the content and organisation of students' writing: EFL teachers' perspective. *Cogent Education*, *10*(2), 2236469. DOI: 10.1080/2331186X.2023.2236469

Mazur, E. (2009). Confessions of a converted lecturer. Retrieved on March 22, 2014 from YouTube: http://www.youtube.com/watch?v=WwslBPj8GgI

McLeod, S. A. (2010). *Simply Psychology; Zone of Proximal Development.* Retrieved May 1, 2014 from http://www.simplypsychology.org/Zone-of- Proximal-Development.html

Musallam, R. (2010). *The effects of using screencasting as a multimedia pre-training tool to manage the intrinsic cognitive load of chemical equilibrium instruction for advanced high school chemistry students*. University of San Francisco.

O'Donoghue, Y. & Clarke, R. (2010). *The flipped classroom model for college algebra: Effects on student achievement*. Doctoral thesis, Colorado State University, Fort Collins, Colorado).

Paesani, K. (2005). "Literary Text and Grammar Instruction: Revisiting the Inductive Presentation". *Foreign Language Annals*, v38 n1 p15-24 Spr 2005, EJ741802. Retrieved on July 17, 2015 from: http://www.eric.ed.gov/

Piaget, J. (1968) *On the development of memory and identity*. Barre, MA: Clark University Press with Barre Publishers.

Pink, D. (2010). Think Tank: Flip-thinking-the new buzz word sweeping the US: the *Telegraph*, *17*.

Qeshta, A.(2016).*The Impact of Using the Flipped Learning Strategy on Concept Development and Reflective Thinking in Biology among Female Tenth Graders.* (Unpublished Master Dissertation), The Islamic University, Gaza Strip, Palestine.

Sainani, K. (2013). flipped classroom Overview. *Retrieved May 16*, 2015.

Sams, A. (2011a). There is No Such Thing as the flipped class: The Flip is in Flux. Retrieved from chemicalsams: http://chemicalsams.blogspot.ca/2014/10/there-isno-such-thing-as-flipped-class.html

Sams, A. (2011b). The flipped class: Shedding light on the confusion, critique, and hype.

The Daily Rift. Retrieved July 15, 2017, from http://www.thedailyriff.com/articles/the-flipped-class-shedding-light-on-the-confusion-critique-and-hype-801.php.

Shehadeh, A. (2022). Classifying K12 blended learning. Mountain View, CA: Insight Institute, Inc.

Soleimani, E. (2019). Flipped vs. unplugged instruction: Sailing EFL learners' oral proficiency through virtual and real learning communities. *International Journal of Instruction, 12*(3), 459-480.

Teng, M. F. (2024). "ChatGPT is the companion, not enemies": EFL learners' perceptions and experiences in using ChatGPT for feedback in writing. *Computers and Education: Artificial Intelligence*, *7*, 100270. ttps://doi.org/10.1016/j.caeai.2024.100270

Tucker, B. (2012). The flipped classroom. *Education next*, *12*(1), 82-83.

Ur, P. (1991). *A Course in Language Teaching; Practice and Theory*. Cambridge University Press.

Uzunboylu, H., & Karagozlu, D. (2015). Flipped classroom: A review of recent literature. *World Journal on Educational Technology, 7*(2), 142-147.

Vadivel, B., Shaban, A. A., Ahmed, Z. A., & Saravanan, B. (2023). Unlocking English Proficiency: Assessing the Influence of AI-Powered Language Learning Apps on Young Learners' Language Acquisition. International Journal of English Language, Education and Literature Studies, 2(6). DOI: https://dx.doi.org/10.22161/ijeel.2.6.7

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Werdiningsih, I., Marzuki, & Rusdin, D. (2024). Balancing AI and authenticity: EFL students' experiences with ChatGPT in academic writing. *Cogent Arts & Humanities*, *11*(1), 2392388. https://doi.org/10.1080/23311983.2024.2392388

Williams, C. (2012). Flipped class method gaining ground. District Administration, 48(1), 64.

Yıldız, Y. & Otacıoğlu, S. G., A. (2017). The Effects of Flipped Learning Model on Student Success in Flute Education. *Route Educational and Social Science Journal*, *4* (6), 254-270.

Zhang, Z. V. (2020). Engaging with automated writing accuracy evaluation (AWE) feedback on L2 writing accuracy: Student perceptions and revisions. *Assessing Writing accuracy, 43*, 100439.

Zhonggen, Y. & Guifang, W. (2016). Academic Achievements and Satisfaction of the Clicker-Aided Flipped Business English Writing accuracy Class. *Journal of Educational Technology & Society,* 19 (2), 298-3.