Castledown

OPEN ACCESS

Technology in Language Teaching & Learning
ISSN 2652-1687
https://www.castledown.com/journals/tltl/

Technology in Language Teaching & Learning, 7(1), 102479 (2025)
https://doi.org/10.29140/tltl.v7n1.102479

Exploring the Role of Digital Writing Tools in Enhancing Academic Writing Skills: The Case of the Hemingway App in EFL Classrooms

MOHSEN ASHRAFGANJOE

Department of Humanities, National University of Skills (NUS), Tehran, Iran
mohsenashrafganjoe@gmail.com

Abstract

This study explored the application and effectiveness of Hemingway-based feedback in enhancing the academic writing skills and overall writing quality of Iranian EFL learners. Given the limited research on integrating digital tools into diverse educational contexts, the study aimed to: (1) examine the impact of the Hemingway app on EFL learners' writing skills and quality; (2) investigate the retention of acquired knowledge; and (3) evaluate its impact on learners' self-efficacy, motivation, and writing anxiety. Two intact classes comprising 40 intermediate TEFL and English Translation students (aged 18–22) from Kerman Islamic Azad University were selected for the study. One class (n = 20) was randomly assigned to the experimental group, which received instruction using the Hemingway app, while the control group (n = 20) followed traditional instruction methods. Both groups completed a writing pre-test and post-test. The study employed a Split-Plot ANOVA (SPANOVA) to assess the intervention's impact on writing performance across three time points. Additionally, the findings revealed that both methods significantly improved participants' writing skills; however, the Hemingway app had a greater impact on content, communicative achievement, organization, and language. While the experimental group outperformed the control group, a slight decline in scores over time was observed. The Hemingway application has significantly enhanced students' self-efficacy in writing by boosting their confidence in generating ideas and using appropriate language, while also reducing their writing anxiety levels. These findings underscore the pedagogical potential of integrating digital tools like the Hemingway app into EFL writing instruction.

Keywords: Hemingway App, traditional feedback, self-efficacy, motivation, writing anxiety, writing skills, writing quality

Introduction

Mastering a foreign language has become an essential requirement in contemporary society for both social and educational purposes. Educational objectives have undergone a remarkable transformation in the modern era, particularly with the integration of educational technologies. This era is characterized by rapid technological advancements, including artificial intelligence, the Internet of Things (IoT), big data, and cloud computing (H. L. Jacobs, 1981; Zhenpeng, 2024). In the 1980s, technology in language classrooms included film, radio, TV, language labs, computers, and interactive video (Cunningham, 1998). Computer-assisted language learning (CALL) also gained prominence during this period (Iandoli, 1990). Although innovative software like MacLang was used (Frommer, 1989), most CALL methods were limited to drill-and-practice activities. However, advances in technology have facilitated the development of more interactive CALL systems and the integration of multimedia tools (Pusack & Otto, 1990). The implementation of digital technology in educational settings is influenced by factors such as community characteristics, socioeconomic status, the readiness and willingness of educators, educational attainment, and the income level of countries. Outside of the most technologically advanced nations, the utilization of computers and devices in classrooms remains limited (Antoninis et al., 2023). Training educators in the effective integration of technology within the classroom is essential to ensure they have unrestricted access to teaching resources and can continuously develop their professional skills. As noted by(Ngo, Chen, & Lai, 2024), Teaching and learning in English as a Second Language (ESL) classrooms often present unique challenges that require innovative solutions to meet the diverse needs of learners, particularly given that writing serves as a critical component of language acquisition (Biju & Vijayakumar, 2023). Writing, as a fundamental tool for communication, serves as a conduit through which students can demonstrate their comprehension and express their cognitive constructs (Moses & Mohamad, 2019). Furthermore, Zamel (1982) viewed writing as a process through which meaning is constructed and conveyed. It is an intricate skill that demands a comprehensive understanding of various components, including spelling, grammar, sentence structure, vocabulary, and textual organization (Javed et al., 2013). Hadi et al., (2021) highlighted the complexity of writing, emphasizing the importance of mastering these fundamental elements to achieve effective written communication. Consequently, teaching writing is considered one of the most challenging skills due to the wide range of aspects involved, including linguistic knowledge, cognitive strategies, self-efficacy, writing anxiety, and motivation (Tangpermpoon, 2008). These factors not only affect the writing process but also pose challenges for language instructors in helping learners develop proficiency in writing. (W. Wu & Wu, 2008). However, if these affective factors are properly addressed, they can significantly enhance the writing process, making it more manageable and effective for learners (Farzana, 2015). Furthermore, integrating technological tools into writing instruction has been identified as a promising approach to address these challenges, as technology can help mitigate affective barriers and foster learner engagement (Mujiyanto & Rukmini, 2018). Several researchers (Mujiyanto & Rukmini, 2018) have highlighted the significance of technology as a crucial tool in language teaching. Technology has been proven to boost student motivation and cultivate essential skills such as creative thinking, effective communication, critical reasoning, productivity, and digital literacy (Tinio, 2002). These skills are particularly important for language development, especially in improving writing abilities (Astuti, Aziz, Sumarti, & Bharati, 2019). Building on these insights, the present study aims to explore the impact of computer-based feedback through Hemingway on Iranian EFL learners' writing skills and writing quality**.**

Literature Review

Impact of Technology on Improving Writing Proficiency in EFL Learners

The process of writing has become increasingly supported by digital technology, which enables students to communicate more easily with others worldwide. The use of Information and

Communication Technology (ICT) in writing instruction offers several advantages, including capturing students' attention, enhancing the learning process, improving vocabulary acquisition, and fostering meaningful learning experiences (Yunus et al., 2013). Technology is an important language teaching tool in this sense, as the use of technology can increase students' motivation and writing proficiencies (Wiyaka et al., 2018). Students' writing ability can be improved by fostering their interest, motivation, and enjoyment of writing through technology (Graham & Perin, 2007). Moreover, technology can also influence writing quality of EFL learners. De Beaugrande and Dressler (1981) identified seven fundamental features that characterize the quality of written discourse: cohesion, coherence, intentionality, acceptability, informativeness, situational relevance, and intertextuality. These features collectively determine the excellence of a text, contributing to its classification as high-quality. McNamara et al., (2010) believed that certain critical features, such as complex sentence structures, a varied vocabulary, and infrequent word usage, can significantly impact writing quality. They indicated that essays of higher quality are distinguished by linguistic attributes linked to text complexity and advanced language utilization. A study undertaken by Xu et al., (2019) examined the efficacy of educational technology applications in enhancing the writing quality of adult English language learners (ELLs). The researchers explored the potential of technology to augment the writing quality of adult ELLs, and their findings showed a meaningful positive relationship based on statistical analysis between technology integration and writing quality improvement.

However, the profound influence of technology on various aspects of writing such content, communicative proficiency, structural organization, and language can be meticulously studied to comprehend the intricate processes involved in writing. Such research can provide valuable insights to enhance language teaching methodologies, assessment techniques, and the overall evolution of language (Faiz et al., 2023). In this regard, content assessment evaluates the degree to which a learner has successfully completed an assigned task, while communicative achievement assesses the appropriateness and effectiveness of the participant's communication style for the given task, including their use of the correct register. Organization refers to the logical structure and orderliness of the learner's written work. Language, on the other hand, concentrates on the breadth and accuracy of vocabulary and grammar usage. These components are pivotal in assessing students' writing competency and fostering holistic language skills development (Canale & Swain, 1980). The emphasis on individual skill components should not be misconstrued as a call for the solitary instruction of such skills, neglecting the broader writing process. Rather, these components should be viewed as interdependent and evolving skills that complement and influence the writing process and its outcome (Khazrouni, 2019). A non-randomized examination conducted by Nguyễn and Baker (2023) to investigate the impact of extensive reading (ER) on the second language writing proficiency of young English language learners. The findings revealed that the experimental group (EG) achieved significantly higher total band scores and subscale scores in content, organization, and language compared to the control group (CG). Notably, no significant differences were observed between the EG and CG in terms of communicative achievement, indicating that ER may positively influence young learners' writing proficiency in specific domains.

However, researchers should strive to ignite students' curiosity and foster their engagement with the material, especially when it comes to expressing their thoughts in written English. In this regard, introducing effective pedagogical tools for writing in English as a Foreign Language (EFL) classrooms is crucial for enhancing students' motivation and writing proficiency (Ashrafganjoe et al., 2022). Within this context, a writing application like Hemingway emerges as a promising solution to address these educational challenges and support the development of writing skills (Imran, 2022).

Hemingway App

The Hemingway App is an online writing tool that aims to enhance the clarity and simplicity of written content. It includes three levels of phrase complexity, including technical, accessible, and default This study specifically focused on the technological level. The app provides color-coded feedback on reading and vocabulary choices, along with voice analysis directly within the text. The text is highlighted with different colors, which correspond to the color-coded feedback on the right side of the screen. Each color represents a distinct feedback category within the Hemingway App: red and yellow highlights indicating wordy sentences, blue highlights indicating adverbs, purple highlights indicating complex phrases, and green highlights indicating redundant passive voice. The app emphasizes eliminating redundant passive voice, reducing adverb usage, simplifying vocabulary, and resolving hard-to-read sentences. It is compatible with cellphones, Chromebooks, and laptops, making it a versatile tool for writers seeking to enhance readability and engagement. Due to its easy-to-use design and comprehensive analysis, the Hemingway App has become an essential resource for writers of all levels, offering practical advice on how to improve the overall quality and impact of their written work. While Sánchez Calderón and da Cunha (2023) study showed that the Hemingway tool was found to be particularly helpful for EFL learners in improving their English writing skills, Jiang and Eslami (2022) showed that the Hemingway may not significantly enhance accuracy, complexity, or meet the needs of all EFL learners. Writing-aided interventions have shown a favorable impact on the learning process. For example, Imran (2022) carried out a study on applying the Hemingway app to enhance students' writing skills. The findings revealed that the Hemingway App significantly improved students' writing skills, demonstrating a notable difference between their writing abilities before and after the intervention.

Interplay of Motivation, Self-Efficacy, and Anxiety in the Writing Process

In addition to the required linguistic acumen and cognitive strategies essential for the act of writing, the complexities of this generative skill are further compounded by factors such as self-efficacy, writing apprehension, and motivational dimensions (W. Wu & Wu, 2008). Krashen (2009) identified motivation, self-efficacy, and anxiety as key emotional factors that play a significant role in second-language learning.

A substantial body of research underscores the negative impact of anxiety on academic achievement (Horwitz et al., 1986; Jebreil et al., 2015; Zrekat et al., 2016). Writing anxiety, in particular, has been shown to hinder the writing performance of first language (L1) learners (Daly & Miller, 1975; Leki, 1990; Piniel, 2013). Kirmizi and Kirmizi (2015), for example, emphasized that "anxiety fosters demotivation and discouragement among students, potentially leading to the development of negative self-efficacy beliefs and attitudes towards writing" (p. 57). Rezaei et al., (2014) examined writing anxiety in a mixed method study on Iranian EFL learners and they found that the most common kind of anxiety is somatic, subsequently accompanied by cognitive anxiety and behaviors aimed at avoidance. Therefore, it is predictable that EFL learners produce a level of anxiety during their engagement in writing activities. Furthermore, anxiety is closely linked to other factors such as self-efficacy, motivation, and writing attitudes.

Motivation and self-efficacy are recognized as additional key factors influencing EFL writing performance, not only in the specific context of writing tasks but also in broader English language learning within the EFL environment (Brown, 2007; Du, 2009; Yusuf, 2011; Zhang et al., 2015). Achievement motivation is defined as an individual's intrinsic desire to perform well and experience a sense of personal accomplishment (Elias et al., 2010). A substantial body of research has indicated a positive correlation between achievement motivation, self-efficacy, and writing outcomes within

academic settings (Pajares & Johnson, 1994; Piniel, 2013). Furthermore, X. Wu et al., (2022) argued that the integration of motivation, self-efficacy, and anxiety related to the use of English plays a crucial role in the acquisition of a second language. The synergy of these factors notably predicts English proficiency among students.

Self-efficacy refers to an individual's belief in their capability to perform successfully, acting as a significant motivator of human behavior. High self-efficacy is associated with increased perseverance and resilience when faced with difficulties during writing tasks (Bandura et al., 1999). Shah et al., (2011) investigated the self-efficacy levels of 120 Malaysian ESL learners, revealing a moderate correlation between their self-efficacy and writing performance. These critical elements contribute significantly to either facilitating or hindering the reception of language input, thereby shaping a learner's likelihood of success in mastering a new language. Sabti et al., (2019) conducted a study on the impact of writing anxiety, writing achievement motivation, and writing self-efficacy on writing performance. The results showed that increased levels of writing anxiety correlated with poorer writing performance, while greater writing self-efficacy and higher achievement motivation were linked to improved writing performance.

Due to the intricate nature of the writing process, students often encounter various challenges within educational settings. Initially, students tend to underestimate the importance of writing due to a lack of emphasis placed on its significance by instructors, who merely assign writing tasks without providing a rationale (Aunurrahman et al., 2017). Secondly, a significant issue arises from many students' inadequate grasp of grammar. Proficiency in grammar is crucial in crafting coherent sentences and producing high-quality written material. Thirdly, students often face challenges related to limited vocabulary, which negatively impacts the quality of their writing. A strong vocabulary is pivotal in helping students articulate their ideas effectively and enhancing the overall quality of their written work (Hadi et al., 2021). Insufficient knowledge of grammar and writing skills can undermine learners' confidence in expressing their ideas, despite their mastery of syntactic, lexical, and grammatical elements (Rico, 2014). Additionally, teaching strategies that do not align with students' learning styles and cultural backgrounds further diminish their confidence and motivation (Ahmad et al., 2013).

As the review of the literature revealed, automated feedback provided by technological tools for improving writing skills is an effective strategy for developing learning in EFL contexts. However, some writing-aided tools, such as Grammarly and Canva App, have been frequently studied (Ashrafganjoe et al., 2022; Masturoh et al., 2023; Yundayani et al., 2019). Furthermore, the aforementioned studies primarily focused on the impact of technology on writing improvement through pre-test and post-test assessments but did not address the effect of technology on the long-term retention of acquired knowledge by language learners. Truscott (2007) highlighted various studies that questioned the effectiveness of grammar correction, pointing to issues such as the lack of control groups, absence of delayed post-tests, and reliance solely on grammar exercises as writing tasks. Additionally, Iran has now reached a point where the utilization of computer-generated feedback to enhance writing proficiency is viable. Nonetheless, existing research indicates that the feedback offered by Hemingway has been underexplored within the Iranian context. In fact, the focus on tools like the Hemingway App is lacking in Iran. Moreover, to date, there has been virtually no study on Iranian EFL learners that investigates the influence of computer-generated feedback by Hemingway on language learners' attitudes.

In light of the problems discussed above, on the one hand, this research aimed to assess the long-lasting effects of the Hemingway App on academic writing skills such as content, communicative achievement, organization, and language. To this end, a delayed post-test was conducted one month after the immediate post-test, following the same format as the initial pre-tests and immediate post-tests.

Additionally, an analysis was conducted on the quality of assignment generated by students in each group to determine which teaching method facilitated more fluent and accurate elaboration of their ideas. On the other hand, this study explored the extent to which exposure to the Hemingway App impacts students' writing anxiety, motivation, and self-efficacy. These factors have consequently garnered the researcher's interest in adopting Hemingway as a writing-aided intervention for instructional delivery. In this connection, the purpose of this study was to probe the impact of Hemingway's automated feedback on the writing skills and attitudes of English as a Foreign Language (EFL) learners. To achieve this, a nonrandomized control group pretest-posttest design was used to collect and analyze data. As a result, based on the identified problems and goals, the research questions outlined below were proposed.

1. Is there any difference between the traditional approach and computer-based instruction on the development of EFL learners' writing skills (content, communicative achievement, organization and language) from pre-tests to post-tests?
2. Is there a significant difference in the retention of learned knowledge between the experimental group and control group from the post-test to the delayed post-test?
3. Which approach of teaching writing is more effective at improving EFL learners' writing abilities in terms of quality?
4. To what extent does the use of the Hemingway App affect students' self-efficacy, writing anxiety and motivation?

Methodology

Research Design

The research used a quasi-experimental design, specifically focusing on a nonequivalent pretest-posttest control group design. This methodology necessitates the formation of two groups: a treatment group, which received a pre-test followed by treatment using the Hemingway App to assess its impact on EFL learners' L2 writing skills (including content, communicative achievement, organization, and language), and a control group that only underwent pre-test and post-test assessments without any intervention (McMillan & Schumacher, 2010). Due to the challenges of random assignment, the study utilized intact, pre-existing groups, administering the pre-test, applying the intervention to the experimental group, and subsequently conducting the post-test. Additionally, the qualitative phase of the study adopted a non-experimental design to gather participants' perceptions regarding self-efficacy, writing anxiety, and motivation through a questionnaire.

Instruments

Diagnostic Language Assessment (DIALANG)

It is a highly respected online diagnostic assessment platform that evaluates second language learners' grammatical proficiency by aligning their performance with the Common European Framework of Reference for Languages (CEFR) (Huhta et al., 2002).

Pre-Test, Post-Test and Delayed Post-Test

A pre-test was administered to evaluate participants' English writing proficiency prior to treatment. Following this, a post-test assessed their learning achievements, and a delayed post-test measured retention of writing knowledge. The topics for the pre-test, immediate post-test, and delayed post-test were selected from the Test of Written English (TWE) series available on the Educational Testing

Service (ETS) website. To ensure validity, these topics were reviewed by three TEFL PhD holders from Islamic Azad University and Shahid Bahonar University in Kerman. Consistency in topics and instructors was maintained across both experimental and control groups. The writing assessments were scored using a rubric based on The Preliminary English Test (PET), (CambridgeAssessmentEnglish, 2020) (see Appendix A). Reliability was evaluated using Cohen's Kappa, with values of .61–.80 indicating good reliability and .81–1 indicating excellent reliability. The pre-test and post-test demonstrated adequate reliability indices as shown in Table 1.

Table 1 *Interrater Reliability of Two Raters in Pre-Test and Post-Test for Writing Skills*

Time	Measure	Value	Confidence Interval (95%)
Pre-test	Cohen's Kappa	.794	(0.84–0.92)
Post-test	Cohen's Kappa	.940	(0.60–1)

Writing Attitude Questionnaires

The subsequent instrument employed in this quantitative study was a questionnaire comprising four sections, which was adapted with minor modifications from research by Le et al., (2023). This questionnaire utilized a five-point Likert scale, ranging from strongly disagree (1) to strongly agree (5), and consisted of 19 items: 5 on self-efficacy, 6 on anxiety, 6 on motivation, and 2 on general perceptions of the Hemingway App. It aimed to investigate differences in perspectives between participants in the experimental group across pre-test and post-test (see Appendix B). To ensure validity, two professors reviewed the questionnaire, providing feedback that led to necessary adjustments for clarity. Revisions included adding "Hemingway App" to relevant items. The questionnaire was developed in English and not translated into Persian, as all participants had sufficient comprehension of the language. Reliability was evaluated through Cronbach's alpha, yielding coefficients of .942 for the pre-test and .952 for the post-test (see Table 2). According to (Keser et al., 2013) internal reliability coefficients around 0.70 are considered adequate, those around 0.80 are very good, and values around 0.90 are excellent. Thus, the questionnaire demonstrated excellent reliability (see Appendix C).

Table 2 *Cronbach's Alpha Reliability Statistics of Writing Attitude Questionnaire*

Time	Cronbach's Alpha	Number of Items
Pre-test	.942	19
Post-test	.952	19

Writing Quality Rubric

During the semester, the instructor provided feedback, comments, and suggestions to both groups of students on various aspects. Additionally, the quality of the writing was evaluated using (H. Jacobs et al., 1981) the ESL composition profile, which comprises five components: content (30%), organization (20%), vocabulary (20%), language use (25%), and mechanics (5%). As a result, each writing sample was evaluated and given a score out of 100 according to these criteria (see Appendix B). Cohen's kappa was used as a way to estimate the inter-rater reliability tests in pre-test and post-test. (see Table 3). As noted by et al., (2015) a Cohen's Kappa of .61–.80 is a good reliability index and .8–1is an excellent reliability index. Based on these criteria this test enjoyed an excellent reliability index.

Table 3 *Interrater Reliability of Two Raters in Pre-Test and Post-Test for Writing Quality*

Time	Measure	Value	Confidence Interval (95%)
Pre-test	Cohen's Kappa	.893	(0.84–0.92)
Post-test	Cohen's Kappa	.813	(0.60–1)

Context and Participants

This study selected a sample of 40 EFL learners (18–22 years old) from a larger pool of 80 using convenience sampling (Dornyei, 2007).

Two intact classes majoring in Teaching English as a Foreign Language (TEFL) and English Translation Studies at Kerman Islamic Azad University were chosen, with one class randomly assigned as the experimental group and the other as the control group. Each class comprised 20 learners, totaling 18 male and 22 female participants. Participants' proficiency levels were assessed using DIALANG, revealing that both groups had writing skills at the B1 level. Notably, none of the learners had prior experience in technology-based classrooms. The first author served as both the researcher and course instructor for the two groups, bringing extensive experience as an EFL instructor and researcher.

Data Collection and Analysis

The data collection process comprised three assessment phases: a pre-test, an immediate post-test, and a delayed post-test, conducted during the spring semester of the 2023 academic year (September to November). Initially, participants completed a writing task as a pre-test to evaluate their English writing proficiency, which required them to produce a text of 150–190 words within a 20-minute timeframe. During the treatment phase, the control group engaged in the composition of eight essays across eight sessions, each adhering to the same word count and time constraints. Following each writing task, participants received feedback from the instructor on key dimensions: content, communicative achievement, organization, and language. Errors were marked for clarity, with corrections provided in the margins. Conversely, the experimental group followed a similar structure but incorporated the Hemingway application. Participants were trained on its functionalities and utilized it to receive automated feedback emphasizing critical writing components. Submissions were made via the Hemingway app to obtain immediate corrective insights, after which hard copies were provided to the instructor for further evaluation. Upon completion of the eight writing sessions, participants undertook an immediate post-test to assess improvements in their writing skills. To evaluate the durability of these improvements, a delayed post-test was administered one month later. This design aimed to discern treatment effects from potential confounding variables, thereby enhancing the study's validity. Writing tasks were assessed using the PET Cambridge writing scale, which encompasses four criteria—content, communicative achievement, organization, and language proficiency—each scored out of five points for a total of 20 points. Two independent raters evaluated each submission, with inter-rater reliability measured using Cohen's Kappa. To address the first research question regarding differences between teaching methodologies, both independent and paired sample t-tests were employed. For the second question concerning long-term effects of the Hemingway application on writing skills, Split-plot ANOVA was utilized. The ESL composition profile created by H. J. Jacobs et al. (1981), which allocates scores based on content (30%), organization (20%), vocabulary (20%), language use (25%), and mechanics (5%), culminating in a total score out of 100. Finally, to explore participants' attitudes towards the Hemingway application in the experimental group, a 5-point Likert scale questionnaire comprising 19 items was administered. This instrument measured dimensions such as self-efficacy, anxiety, motivation, and overall perceptions regarding the application's effectiveness in enhancing writing skills.

Results

To address the first research question, two paired-samples t-tests were conducted to determine whether participants showed statistically significant improvement in writing skills from pre-test to post-test. Prior to analyzing the results, a Shapiro-Wilk test was applied to confirm the assumption of normality for both pre-test and post-test distributions. The test indicated that the assumption of normality was retained, as the p-value exceeded the 0.05 threshold (see Table 4).

Table 4 *Normality Assumption for Pre-Test and Post-Test of Two Groups*

	Shapiro–Wilk		
Time	**Statistic**	**df**	**Sig.**
Control. experimental. pre-test	.966	40	.273
Control. experimental. post-test	.956	40	.119

To compare the mean scores of the four specific features between the two groups before and after the study, two independent samples t-tests were conducted (see Table 5). The results indicated no significant differences in the pre-test scores for each feature between the groups ($p > 0.05$).

Table 5 *Comparison for the Four Specific Features between the Two Groups before the Study*

Pre-test	Features	Group	N	Mean	SD	Mean Difference	t	df	Sig.
Equal variances assumed	Content	Control	20	1.35	.933	–.100	–.310	38	.758
		Experimental	20	1.45	1.099				
Equal variances assumed	Communicative achievement	Control	20	1.95	.759	.050	.196	38	.846
		Experimental	20	1.90	.852				
Equal variances assumed	Organization	Control	20	1.75	.851	–.100	–.366	38	.716
		Experimental	20	1.85	.875				
Equal variances assumed	Language use	Control	20	1.80	1.056	.200	.602	38	.551
		Experimental	20	1.60	1.046				

Table 6 presents the data comparing the mean scores of the four features between the two groups in the post-test. The results revealed that there was no difference in three features (content, organization, language use) in the post-test between the two groups. But the EFL learners' writing performance in the experimental group ($M = 4.45$, $SD = .686$) in terms of *communicative achievement* was much better than that in the control group ($M = 3.45$, $SD = .686$); $t(38) = – 4.607$, $p < 0.05$.

As illustrated in Table 7, the results of the paired samples t-test showed a statistically significant improvement in writing skills for both the control group ($t(19) = -13.582$, $p < 0.05$, $d = –3.037$) and the experimental group ($t(19) = –17.027$, $p < 0.05$, $d = –3.807$), with both groups exhibiting large effect sizes that underscore the substantial improvements from pre-test to post-test scores.

As Table 8 shows, the independent samples t-test was run with equal variances assumed, revealing no statistically significant difference in writing skills between the experimental and control groups on pre-test ($t(38) = .081$, $p > 0.05$ (two-tailed). In contrast, post-test results showed a statistically significant

Table 6 *Comparison for the Four Specific Features between the Two Groups on Post-Test*

Post-test	Features	Group	N	Mean	SD	Mean Difference	t	df	Sig
Equal variances assumed	Content	Control	20	3.30	.865	–.200	–.721	38	.475
		Experimental	20	3.50	.889				
Equal variances assumed	Communicative achievement	Control	20	3.45	.686	–1.000	–4.607	38	.000
		Experimental	20	4.45	.686				
Equal variances assumed	Organization	Control	20	3.95	.945	–.100	–.399	38	-.608
		Experimental	20	4.05	.605				
Equal variances not assumed	Language use	Control	20	3.80	1.508	–.600	–1.622	26.432	.117
		Experimental	20	4.40	.681				

Table 7 *Paired-Sample T-Tests Comparing Means of Control Group in Pre-Test and Post-Test*

Control Group	N	Mean	Std. Deviation	t	*P*	*df*	*d*
Pre-test	20	6.85	1.755				
Pair				–13.582	.000	19	–3.037
Post-test	20	14.5	2.039				
Experimental Group							
Pre-test	20	6.80	2.142				
Pair				–17.027	.000	19	–3.807
Post-test	20	16.40	1.698				

Table 8 *Independent-Samples T-Test Comparing Means of Experimental and Control Groups in Pre-Test and Post-Test*

Time	Levene's Test for Equality of Variances			t-test for Equality of Means			
		F	Sig.	t	df	Sig. (2-tailed)	*d*
Pre-test	Equal variances assumed	1.868	.180	.081	38	.936	.023
Post-test	Equal variances assumed	.230	.634	–3.202	38	.003	–.777

difference in writing skills between the groups, with the experimental group outperforming the control group ($t(38) = –3.202$, $p < 0.05$ (two-tailed) $d = –.777$ representing large effect size). The results suggested that the intervention applied to the experimental group led to a meaningful improvement in post-test performance compared to the control group.

However, researchers may sometimes wish to combine both approaches in a single study, using one independent variable as between-subjects (groups) and another as within-subjects (times). These hypotheses can be tested with Split-plot ANOVA (SPANOVA), also known as Mixed between-within subjects ANOVA.Table 9 shows the results of descriptive statistics among pretest, post-test and delayed post-test between the control and the experimental groups. According to Table 9, the control group had

a slight increased in the mean score in pretest (M = 6.85, SD = 1.75), but the experimental group had the higher mean scores from post-test (M = 16.4, SD = 1.69) to delayed post-test (M = 15.2, SD = 1.76).

Table 9 *Descriptive Statistics of Pre-Test, Post-Test and Delayed Post-Test Scores*

	Groups	Mean	Std. Deviation	N
Pre-test	Control	6.8500	1.75544	20
	Experimental	6.8000	2.14231	20
	Total	6.8250	1.93334	40
Post-test	Control	14.5000	2.03909	20
	Experimental	16.4000	1.69830	20
	Total	15.4500	2.08720	40
Delayed post-test	Control	11.2000	1.54238	20
	Experimental	15.2000	1.76516	20
	Total	13.2000	2.60374	40

The mean differences among the pre-test, post-test, and delayed post-test scores were substantial. However, the Multivariate Test table from the SPANOVA output determined whether these differences were statistically significant. Like many parametric tests, SPANOVA has specific assumptions that must be satisfied prior to analysis. One key assumption is the equality of covariance matrices, assessed using Box's M test; for this assumption to hold, the p-value must exceed 0.001. As shown in Table 10, the p-value for this test is greater than this threshold, indicating that the assumption was not violated.

Table 10 *Test of Equality of Covariance Matrices*

Box's M	1.968
F	0.300
df1	6
df2	10462.189
Sig.	.937

Another assumption of SPANOVA is the equality of error variances, or homogeneity of groups, which is typically assessed using Levene's Test of Equality of Error Variances. Non-significant values are required for the pre-test, post-test, and delayed post-test scores to satisfy this assumption. Table 11 shows that the p-values for the pre-test, post-test and delayed post-test scores are .180, .634 and .369 respectively, that is, they are non-significant (p >.05).

Table 11 *Levene's Test of Equality of Error Variances*

	Levene Statistic	df1	df2	Sig.
Pre-test	1.863	1	38	.180
Post-test	.230	1	38	.634
Delayed post-test	.827	1	38	. 369

After ensuring non-violation of assumptions, the SPANOVA was conducted, and the results are presented in Table 12. It is important to note that all these tests serve the same purpose, so it is advisable to focus on just one, preferably Pillai's Trace or Wilks' Lambda, as these two are more effective in revealing interactions. As presented in Table 12, the results of the multivariate tests indicated significant interactions between time and group in the dataset, evidenced by an interaction effect of (F(2,37) = 15.764, $p < 0.05$ $\eta 2$ =.460 representing a small effect size). Furthermore, the main effect of time is also significant, (F(2,37) = 283.239, $p < 0.05$ $\eta 2$ = .939 representing a large effect size). The results of SPANOVA analysis indicated a significant difference between the two groups across the three assessed time points, underscoring the importance of both the interaction effects and the main effects in this context.

Table 12 *Tests of Main and Interaction Effects*

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Time	Pillai's Trace	.939	283.239	2.000	37.000	.000	.939
	Wilks' Lambda	.061	283.239	2.000	37.000	.000	.939
	Hotelling's Trace	15.310	283.239	2.000	37.000	.000	.939
	Roy's Largest Root	15.310	283.239	2.000	37.000	.000	.939
Time *Group	Pillai's Trace	.460	15.764	2.000	37.000	.000	.460
	Wilks' Lambda	.540	15.764	2.000	37.000	.000	.460
	Hotelling's Trace	.852	15.764	2.000	37.000	.000	.460
	Roy's Largest Root	.852	15.764	2.000	37.000	.000	.460

Table 13 indicates that the main effect of group, had been significant (F (1,38) = .000, $p < 0.05$, $\eta 2$ =.437 representing a small effect size).

Table 13 *Main Effect of Group*

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	16779.675	1	16779.675	4330.239	.000	.991
Group	114.075	1	114.075	29.439	.000	.437
Error	147.250	38	3.875			

The SPANOVA analysis indicated a substantial difference in writing skills between the groups. Specifically, Wilks' Lambda indicated a p-value of .000, which is less than 0.05, demonstrating a significant difference across the pre-test, post-test, and delayed post-test. This suggests that time as a variable significantly influenced L2 writing skills. As shown in Table 12, the interaction of time and group also yielded a p-value of .000, confirming that L2 writing skills were significantly affected by both factors. While SPANOVA indicates differences among groups, it does not specify their source. To pinpoint these differences, pairwise comparisons were conducted.

Table 14 presents the results of these comparisons over the course of the three assessment periods. The control group exhibited substantial performance differences at all three time points. In contrast, the experimental group showed significant differences across most time points; however, the comparison between time 2 and time 3 did not reach statistical significance ($p > 0.05$). Overall, both groups demonstrated meaningful changes over time, except for the non-significant difference between time

2 and time 3 in the experimental group. Additionally, a significant main effect was observed for both the timing of tests and group. Notably, there was a sharp increase in mean scores from the pre-test to post-test for both groups. However, while the control group experienced a substantial decrease from post-test to delayed post-test, the experimental group saw only a modest decline in mean scores during this interval (see Table 9).

Table 14 *Pairwise Comparison between Groups in Three Times*

Group	(I) Time	(J) Time	Mean Difference (I-J)	Sig
Control	1	2 *	-7.650	.000
		3 *	-4.350	.000
	2	1 *	7.650	.000
		3 *	3.300	.000
	3	1 *	4.350	.000
		2 *	-3.300	.000
Experimental	1	2 *	-9600	.000
		3 *	-8.400	.000
	2	1 *	9.600	.000
		3 *	1.200	.160
	3	1 *	8.400	.000
		2 *	-1.200	.160

In order to compare performance in three different times, another pairwise was conducted to compare groups' performance in three times. Table 15 indicates that there was no statistically significant difference found between the treatment and control groups in the pre-test results. (MD = 0.50, p = 0.619, $p > 0.05$). Conversely, statistically significant differences were identified in the post-test (MD = –1.900, $p < 0.05$) and delayed post-test (MD = –4.000, $p < 0.05$), with the experimental group outperforming the control group.

Table 15 *Pairwise Comparison between Times in Two Groups*

Time	(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.[b]
1	Control	Experimental	050	.619	.936
2	Control	Experimental	-1.900*	.593	.003
3	Control	Experimental	-4.000*	.524	.000

Based on estimated marginal means
*. The mean difference is significant at the .05 level.
b. Adjustment for multiple comparisons: Bonferroni.

To address the third research question, two paired-samples t-tests were run to determine whether participants showed statistically significant improvement in writing skills from pre-test to post-test and two independent sample t-tests were run to show the differences between two groups in pre-test and post-test. Prior to analyzing the results, a Shapiro–Wilk test was applied to confirm the assumption of normality for both pre-test and post-test distributions. The test indicated normal distribution in both groups is retained ($p > 0.05$) (see Table 16).

Table 16 *Normality Assumption for Pre-Test and Post-Test of Two Groups*

	Shapiro–Wilk		
	Statistic	**df**	**Sig.**
Control. experimental. pre-test	.954	40	.107
Control. experimental. post-test	.985	40	.875

The results of independent sample t-test on the pre-test revealed no statistically significant difference between the control and experimental groups ($t(38) = 0.73$, $p > 0.05$ (two-tailed) (see Table 17). However, an independent sample t-test conducted at the semester's end revealed that the treatment group significantly outperformed the control group in paragraph writing quality ($t(38) = -2.485$, $p < .05$ (two-tailed), $d = -.777$ representing large effect size) (see Table 17). While both groups showed improvement in writing quality from the pre-test to the post-test, the experimental group's gains were significantly greater. This substantial improvement in the experimental group is likely attributable to the feedback and comments provided on their writing tasks through Hemingway App throughout the semester, which enhanced their performance relative to the control group.

Table 17 *Independent-Samples T-Test for the Quality of Writing Tasks on the Pre-Test and Post-Test Across Control and Experimental Groups*

Time	**Levene's Test for Equality of Variances**			**t-test for Equality of Means**		
		F	**Sig.**	**t**	**df**	**Sig. (2-tailed)**	***d***
Pre-test	Equal variances assumed	.179	.675	.073	38	.942	.023
Post-test	Equal variances assumed	.389	.536	–2.485	38	.019	–.777

Two paired-sample t-tests were conducted to assess the extent of improvement in writing task quality within each group from pre-test to post test (see Table 18). The results indicated significant improvements in writing quality for both the experimental ($t(19)=-14.963$ $p < 0.05$, $d = -4.018$ representing large effect size) and control groups ($t(19)= -17.968$ $p<0.05$, $d = -3.346$ representing large effect size).

Table 18 *Paired-Samples T-Test for the Quality of Writing Tasks Between Control and Experimental Groups*

Control Group	**N**	**Mean**	**Std. Deviation**	**t**	***P***	***df***	***d***
Pre-test	20	54.0500	9.76285				
Pair				–-14.963	.000	19	–3.346
Post-test	20	83.2500	5.50478				
Experimental Group							
Pre-test	20	53.8500	9.76285				
Pair				–17.968	.000	19	–4.018
Post-test	20	87.3000	4.90005				

Table 19 outlines the experimental group's attitudes toward the Hemingway software before and after using the program. The results indicated a noticeable increase in agreement with the software following its implementation. Specifically, the percentage of participants selecting "agree" and "strongly agree" increased from 25% and 35% in the pretest to 85% and 75% in the post-test, respectively. These findings suggested that the observed levels of agreement significantly exceeded initial expectations.

Table 19 *Frequencies and Percentages Attitude towards Hemingway*

Time	Group	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
Pre-test	Experimental group N=20 %	8	14	11	5	7
		40%	70%	55%	25%	35%
Post-test	Experimental group N=20 %	8	7	13	17	15
		40%	35%	65%	85%	75%

The analysis of students' attitudes from the questionnaire responses (Table 20) indicated a generally positive perception of the Hemingway application. In the area of self-efficacy in writing (items 1–5), four out of five items surpassed the midpoint on the Likert scale, with Item 4 receiving the highest positive feedback. Regarding writing anxiety (items 6–11), all items except Item 11 reflected favorable perceptions among participants. Motivation-related items (items 12–17) also showed a positive

Table 20 *Students' Attitudes toward using Hemingway*

Taxonomy of the questionnaire	Items	Responses				
		SD	D	U	A	SA
Self-efficacy	1	12.5	22.5	22.5	25.0	17.5
	2	5.0	25.0	27.5	25.0	17.5
	3	7.5	27.5	25.0	30.0	10.0
	4	5.0	22.5	27.5	32.5	12.5
	5	15.0	25.0	27.5	22.5	10.0
Writing anxiety	6	7.5	30.0	22.5	32.5	7.5
	7	5.0	30.0	27.5	25.0	12.5
	8	10.0	27.5	22.5	27.5	12.5
	9	5.0	30.0	22.5	27.5	15.0
	10	5.0	30.0	29.0	31.0	5.0
	11	20.0	20.0	22.5	27.5	10.0
Motivation	12	15.0	23.0	22.5	19.5	20.0
	13	15.0	25.0	15.0	27.5	17.5
	14	12.5	25.0	22.5	15.0	25.0
	15	20.0	20.0	20.0	25.0	15.0
	16	7.5	30.0	22.5	30.0	10.0
	17	7.5	26.5	30.0	26.0	10.0
The effectiveness General ideas	**18**	**5.0**	**30.0**	**20.0**	**32.5**	**12.5**
Future application	19	10.0	17.5	30.0	17.5	25.0

trend, with Item 13 achieving the highest affirmative responses. Concerning general feedback on the writing practice process with Hemingway, a majority found the tool effective, with 32.5% responding "Agree" and 12.5% "Strongly Agree." While 30% remained neutral ("Undecided"), 27.5% expressed reluctance to continue using Hemingway; however, a notable 42.5% indicated a positive inclination toward its ongoing use. These findings suggest that interest in continuing to use Hemingway for writing practice exceeds the number of students who view the application negatively.

Discussion

The present study aimed to investigate the effects of computer-based feedback via the Hemingway app and teacher-generated feedback on the writing skills of Iranian EFL learners, specifically focusing on content, communicative achievement, language use, and organization. Additionally, the study sought to assess the extent of improvement in writing quality among EFL learners following the utilization of the Hemingway app.

The first research question aimed to determine whether there was a significant difference between the traditional method of teaching writing and Hemingway-based instruction in enhancing writing skills (content, communicative achievement, organization, and language) from pretest to posttest. Based on the results of paired and independent sample t-test, it can be concluded that both the experimental and control groups showed considerable improvement in four writing features over time and even the treatment group demonstrated superior performance in the communicative achievement feature on the post-test. Based on the results, experimental group outperformed control group on post test. Consequently, it can be inferred that Hemingway-based learning is more effective in boostingthe writing skills of EFL learners. regarding the success of the Hemingway-based feedback and teacher feedback, this study draws support from Ashrafganjoe et al. (2022) and Zaini and Mazdayasna (2015) study on the impact of computer-based instruction in improving EFL learners' writing skills. Their findings revealed that both Grammarly® software corrective feedback and teacher corrective feedback are beneficial for enhancing EFL learners' writing abilities. However, Grammarly® software was found to be more effective in developing L2 writing skills in their final analysis. The results of Imran (2022) study showed that Hemingway app can enhance students' writing skill, but these studies contradict those of Omar, Miah, and Belmasrour (2014) in which they believed that computer-based feedback might have affected the writing skills negatively. Furthermore Fazio (2001) and Kepner (1991) argued that teacher-written feedback in L2 writing has been found to be ineffective.

The second research question examined the retention of learned knowledge between the experimental and control groups from the post-test to the delayed post-test. SPANOVA results indicated a significant difference in writing skills across three time points, with both time and group interactions influencing outcomes. Notably, both groups showed a marked increase in mean scores from the pre-test to the post-test. However, the control group experienced a substantial decline in mean scores from the post-test to the delayed post-test, while the experimental group demonstrated only a modest decrease. This suggests that the decay of writing skills was significantly greater in the control group. The observed decline in mean scores for both groups indicated a loss of writing knowledge over time, but the interaction between time and group suggested that this reduction cannot be solely attributed to time alone. These findings align with Sarpong et al., (2024) who reported a similar reduction in learned knowledge between experimental and control groups from post-test to delayed post-test. They posited that conventional classrooms often lack technological resources, hindering students' ability to learn at their own pace and resulting in lower retention. In contrast, this study challenges the conclusions of, as both the experimental and control groups exhibited significant enhancements in their immediate and delayed writing post-tests.

Based on the positive impact of computer-based feedback through Hemingway on the writing quality of EFL learners revealed that both groups showed improvements from pre-test to post-test but the experimental group demonstrating superior performance in writing quality on the post-test. These results align with the conclusions drawn by Imran (2022), who posited that the integrating the Hemingway App significantly enhances the writing skills quality of English as a Foreign Language (EFL) learners. This assertion is further supported by Ablezgova (2018), who noted that the Hemingway App effectively provides feedback on various linguistic features, including grammar aspects such as passive voice and verb tense, as well as spelling, vocabulary choices, organization, structure, and word choice. Conversely, the results of this study diverged from those reported by Ismael et al., (2022), which indicated that automated feedback negatively influenced writing quality among students at two universities in Sulaimaneyah City. Overall, the research presented a nuanced perspective on the influence of automated feedback mechanisms on student writing skills across different educational contexts.

The study probed the effect of computer-based feedback via the Hemingway App on the writing quality of EFL learners. Results indicated that both the experimental and control groups showed improvements from pre-test to post-test, with the experimental group outperforming the control group in writing quality on the post-test. This aligns with Imran (2022), who found that integrating the Hemingway app significantly enhances EFL learners' writing quality. Ablezgova (2018) further supports this by highlighting the app's effectiveness in providing feedback on various linguistic features, including grammar, spelling, vocabulary, organization, and structure. However, this study's findings contrast with those of Ismael et al. (2022), which reported that automated feedback negatively impacted writing quality among students at two universities in Sulaimaneyah City. Overall, this research offers a nuanced perspective on the effects of automated feedback mechanisms on student writing proficiency across diverse educational contexts.

The study examined learners' attitudes towards their L2 writing skills development after using the Hemingway software for immediate corrective feedback. Results from the writing attitude questionnaire indicated a significant positive influence of the Hemingway application on students' self-efficacy in writing. Specifically, the data revealed that Hemingway enhanced students' confidence in generating ideas and employing appropriate language, while also reducing writing anxiety levels. Key factors contributing to this positive perception included proficiency in using the Hemingway tool, frequency of practice, and time pressure. The application played a crucial role in fostering motivation throughout the learning process, with student interests significantly influencing motivation and highlighting the complex interplay between tool usage and learner engagement. This finding is novel in the literature but aligns with Silva Falcón (2023) who argued that automated writing tools like Hemingway not only improve EFL learners' writing skills but also cultivate positive attitudes essential for effective performance. The results reinforced the notion that Automated Writing Evaluation (AWE) can enhance self-efficacy and motivation while potentially alleviating some aspects of writing anxiety (Sari & Han, 2024). Shen and Tao (2024) demonstrated that technology-based writing self-efficacy significantly mitigates writing anxiety among EFL learners. Additionally, it suggested that enhancing metacognitive strategies can effectively reduce writing anxiety by boosting learners' confidence in using technology-driven tools. However, this study diverges from Bandura (1977) assertion regarding self-efficacy, which posits that high self-efficacy does not always yield desired outcomes, especially in anxiety-laden contexts. Learners with high self-efficacy may overestimate their capabilities, leading to insufficient effort and negative experiences that could diminish their self-efficacy. In a recent investigation by Ma, Noordin, and Razali (2022), nearly half of the studies examined indicated a significant reduction in foreign language anxiety due to technology-assisted instruction, while the remainder showed no statistically significant impact.

Conclusion

The findings of this study present substantial evidence of the Hemingway app's effectiveness in enhancing L2 writing skills among Iranian EFL learners, particularly when compared to traditional instructional methods. Specifically, the application of the Hemingway app has led to significant improvements in students' writing quality scores, indicating that learners perform better after utilizing this tool. Furthermore, students have expressed favorable attitudes towards their learning experiences with the app, reporting increased confidence and reduced inhibition in their writing practices. This enhancement in self-efficacy appears to have motivated students further, fostering a more engaged and productive approach to writing. Providing feedback through technology can help educators to concentrate on advanced writing skills, while simultaneously enhancing students' motivation and the overall quality of their writing. The integration of automated feedback tools in language instruction presents a significant opportunity to enhance pedagogical efficiency and effectiveness. Given the inherent challenges and time constraints associated with providing individualized feedback, these tools serve as valuable adjuncts for language instructors, enabling them to address minor errors in student assignments promptly. Consequently, educators can allocate more attention to addressing substantive language issues that require nuanced intervention. Furthermore, the capacity of these tools such as Hemingway app to deliver instantaneous feedback empowers learners, fostering a sense of autonomy and encouraging sustained engagement with their writing tasks. In an era characterized by increasing classroom sizes and heightened demands for personalized support, automated feedback mechanisms emerge as a cost-effective solution to alleviate instructor workload and expedite grading processes. Thus, they not only enhance the quality of written outputs but also reinforce learner motivation, ultimately contributing to improved educational outcomes in language acquisition.

The findings of this study should be generalized cautiously due to several limitations. First, the participants were intermediate EFL learners, indicating a need for replication across diverse proficiency levels. Second, the sample size was limited to 40 participants; thus, larger samples would enhance the validity of the results. While learners' motivation, self-efficacy, and writing anxiety were assessed through questionnaires, incorporating qualitative methods such as interviews or observations could enrich the findings. Additionally, the study's duration of one term (eight sessions) constrained a more in-depth exploration of the impact of Hemingway-based feedback on EFL learners' writing skills and quality. Therefore, future research should investigate the effects of technology-based feedback on EFL learners' academic writing skills over an extended period to better understand the various factors influencing their development.

References

Ablezgova, D. (2018). *Educational Space For English Learning Teenagers.* Paper presented at the ProfMarket: Education. Language. Success.

Ahmad, N., Khan, F. N., & Munir, N. (2013). Factors affecting the learning of English at secondary school level in Khyber Pakhtunkhwa, Pakistan. *International Journal of English Language and Literature Studies, 2*(2), 95–101.

Antoninis, M., Alcott, B., Al Hadheri, S., April, D., Fouad Barakat, B., Barrios Rivera, M., . . . Caro Vasquez, D. (2023). *Global Education Monitoring Report 2023: Technology in education: A tool on whose terms*?

Ashrafganjoe, M., Rezai, M. J., & Elhambakhsh, S. E. (2022). Providing computer-based feedback through Grammarly® in writing classes. *Journal of Language and Translation, 12*(2), 163–176.

Astuti, A. P., Aziz, A., Sumarti, S. S., & Bharati, D. A. L. (2019). *Preparing 21st century teachers: Implementation of 4C character's pre-service teacher through teaching practice.* Paper presented at the Journal of Physics: Conference Series.

Aunurrahman, A., Hamied, F. A., & Emilia, E. (2017). A joint construction practice in an academic writing course in an Indonesian university context. *Celt: A Journal of Culture, English Language Teaching & Literature, 17*(1), 27.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review, 84*(2), 191.

Bandura, A., Freeman, W. H., & Lightsey, R. (1999). *Self-efficacy: The exercise of control.* Springer.

Biju, L. G., & Vijayakumar, S. (2023). Systematic review on the impact of technology-supported writing environment for developing writing skills. *Computer-Assisted Language Learning Electronic Journal, 24*(3), 290–313.

Brown, T. A. (2007). Temporal course and structural relationships among dimensions of temperament and DSM-IV anxiety and mood disorder constructs. *Journal of Abnormal Psychology, 116*(2), 313.

Cambridge Assessment English. (2020). *Assessing writing for Cambridge English Qualifcations: A guide for teachers.*

Canale, M., & Swain, M. (1980). Theoretical bases of com-municative approaches to second language teaching and testing. *Applied linguistics, 1*(1), 1–47.

Daly, J. A., & Miller, M. D. (1975). The empirical development of an instrument to measure writing apprehension. *Research in the Teaching of English, 9*(3), 242–249.

De Beaugrande, R.-A., & Dressler, W. U. (1981). *Introduction to text linguistics* (Vol. 1): Longman London.

Dornyei, Z. (2007). *Research methods in applied linguistics*: Oxford University Press.

Du, X. (2009). The affective filter in second language teaching. *Asian Social Science, 5*(8), 162–165.

Elias, H., Noordin, N., & Mahyuddin, R. H. (2010). Achievement motivation and self-efficacy in relation to adjustment among university students. *Journal of Social Sciences, 6*(3), 333–339.

Faiz, R., Bilal, H. A. B., Asghar, I., & Safdar, A. (2023). *Optimizing ChatGPT as a writing aid for EFL learners: Balancing assistance and skill development in writing proficiency.* Paper presented at the Linguistic Forum-A Journal of Linguistics.

Farzana, S. F. (2015). Affective factors In English language teaching for professionals. *Anglisticum Journal, 1*, 20–35.

Fazio, L. L. (2001). The effect of corrections and commentaries on the journal writing accuracy of minority-and majority-language students. *Journal of Second Language Writing, 10*(4), 235–249.

Gharehblagh, N. M., & Nasri, N. (2020). Developing EFL elementary learners' writing skills through mobile-assisted language learning (MALL). *Teaching English with Technology, 20*(1), 104–121.

Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology, 99*(3), 445.

Hadi, M. S., Izzah, L., & Paulia, Q. (2021). Teaching writing through Canva application. *Journal of Languages and Language Teaching, 9*(2), 228–235.

Haerazi, H., Irawan, L. A., Suadiyatno, T., & Hidayatullah, H. (2020). Triggering preservice teachers' writing skills through genre-based instructional model viewed from creativity. *International Journal of Evaluation and Research in Education, 9*(1), 234–244.

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal, 70*(2), 125–132.

Huhta, A., Luoma, S., Sajavaara, K., & Takala, S. O. M., & Teasdale. A. (2002). *A diagnostic language assessment system for adult learners*.

Iandoli, L. J. (1990). CALL and the profession: The current state. *The French Review, 64*(2), 261 (2002). 272.

Imran, M. C. (2022). Applying Hemingway app to enhance students' writing skill. *EDULEC: Education, Language, and Culture Journal, 2*(2), 180 (2002). 185.

Ismael, K. O., Saeed, K. A., Ibrahim, A. S., & Fatah, D. S. (2022). Effects of auto-correction on students' writing skill at three different universities in Sulaimaneyah City. *Arab World English Journal*.

Jacobs, H. J., Zinkgraf, S. A., Wormuth, D. R., Faye Hartfiel, V., & Hughey, J. B. (1981). *Testing ESL composition: A practical approach.* Rowley, MA: Newbury House.

Jacobs, H. L. (1981). *Testing ESL composition: A practical approach. English composition program*: ERIC.

Javed, M., Juan, W. X., & Nazli, S. (2013). A study of students' assessment in writing skills of the English language. *International Journal of Instruction, 6*(2).

Jebreil, N., Azizifar, A., Gowhary, H., & Jamalinesari, A. (2015). Study on writing anxiety among Iranian EFL students. *International Journal of Applied Linguistics and English Literature, 4*(2), 68–72.

Jiang, W., & Eslami, Z. R. (2022). Effects of computer-mediated collaborative writing on individual EFL writing performance. *Computer Assisted Language Learning, 35*(9), 2701–2730.

Kepner, C. G. (1991). An experiment in the relationship of types of written feedback to the development of second-language writing skills. *The Modern Language Journal, 75*(3), 305–313.

Keser, H., Eşgi, N., Kocadağ, T., & Bulu, Ş. (2013). Validity and reliability study of the internet addiction test. *Mevlana International Journal of Education, 3*(4).

Khazrouni, M. (2019). Assessment for improving ESL learners' writing skills among undergraduate students: A case study of Skyline university college. *International Journal of English Language Teaching, 7*(1), 30–44.

Kirmizi, Ö., & Kirmizi, G. D. (2015). An investigation of L2 learners' writing self-efficacy, writing anxiety and Its causes at higher education in Turkey. *International Journal of Higher Education, 4*(2), 57–66.

Krashen, S. (2009). Principles and practice in second language acquisition.

Le, A. N., Bo, L. K., & Nguyen, N. M. (2023). Canva-based e-portfolio in L2 writing instructions: Investigating the effects and students' attitudes. *Computer Assisted Language Learning, 42*(1), 41–62.

Leki, I. (1990). Potential problems with peer responding in ESL writing classes. *CATESOL journal, 3*(1), 5–19.

Ma, M., Noordin, N., & Razali, A. B. (2022). Effect of educational technology on students' foreign language anxiety: A thematic literature review. *World Journal of English Language, 12*(6).

Masturoh, S., Kusumo, A. P., & Sitoresmi, N. (2023). The use of Canva to improve students' writing skills on procedure text. *Innovative: Journal of Social Science Research, 3*(3), 4681–4690.

McMillan, J. H., & Schumacher, S. (2010). *Research in education: Evidence-based inquiry, MyEducationLab series*. Pearson.

McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). Linguistic features of writing quality. *Written Communication, 27*(1), 57–86.

Mohammadbeigi, A., Mohammadsalehi, N., & Aligol, M. (2015). Validity and reliability of the instruments and types of measurments in health applied researches. *Journal of Rafsanjan University of Medical Sciences, 13*(12), 1153–1170.

Moses, R. N., & Mohamad, M. (2019). Challenges faced by students and teachers on writing skills in ESL Contexts: A literature review. *Creative Education, 10*(13), 3385.

Mujiyanto, J., & Rukmini, D. (2018). Students' Perception on the Usefulness of ICT-Based Language Program. *English Language Teaching, 11*(2), 53–60.

Ngo, T. T.-N., Chen, H. H.-J., & Lai, K. K.-W. (2024). The effectiveness of automated writing evaluation in EFL/ESL writing: A three-level meta-analysis. *Interactive Learning Environments, 32*(2), 727–744.

Nguyễn, M. C., & Baker, J. R. (2023). Exploring the effect of extensive reading on young English learners' second language writing achievement in a Vietnamese primary school setting. *International Journal of Language Education, 7*(4), 673–685.

Omar, A., Miah, M., & Belmasrour, R. (2014). Effects of technology on writing. *International Journal of Science and Applied Information Technology*, *3*(2).

Pajares, F., & Johnson, M. J. (1994). Confidence and competence in writing: The role of self-efficacy, outcome expectancy, and apprehension. *Research in the Teaching of English*, 313–331.

Piniel, K. (2013). L2 motivation, anxiety and self-efficacy: The interrelationship of individual variables in the secondary school context. *Studies in Second Language Learning and Teaching, 3*(4), 523–550.

Pusack, J. P., & Otto, S. K. (1990). Priority: Instruction applying new technologies. *Foreign Language Annals, 23*(5), 409.

Rezaei, M., Jafari, S., & Younas, M. (2014). Iranian EFL students' writing anxiety: levels, causes and implications. *English for Specific Purposes World, 42*(15), 1–10.

Rico, L. J. A. (2014). Identifying factors causing difficulties to productive skills among foreign languages learners. *Opening Writing Doors Journal, 11*(1), 65–86.

Sabti, A. A., Md Rashid, S., Nimehchisalem, V., & Darmi, R. (2019). The Impact of writing anxiety, writing achievement motivation, and writing self-efficacy on writing performance: A correlational study of Iraqi tertiary EFL Learners. *SAGE Open, 9*(4), 2158244019894289.

Sánchez Calderón, S., & da Cunha, I. (2023). Developing autonomy in English writing skills: A study on EFL learners' selection criteria of open access and online self-correction tools. *CIRCULO de Linguistica Aplicada a la Comunicacion, 95*.

Sari, E., & Han, T. (2024). The impact of automated writing evaluation on English as a foreign language learners' writing self-efficacy, self-regulation, anxiety, and performance. *Journal of Computer Assisted Learning*.

Sarpong, V., Appiah-Twumasi, E., Ateko, K. D., & Tanko, S. (2024). Impact of flipped learning pedagogy on senior high school physics students' attainment, retention and feedback in studying electronics. *African Journal of Educational Studies in Mathematics and Sciences, 20*(1), 33–46.

Shah, P. M., Mahmud, W., Din, R., Yusof, A., & Pardi, K. M. (2011). Self-efficacy in the writing of Malaysian ESL learners. *World Applied Sciences Journal, 15*(11), 8–11.

Shen, X., & Tao, Y. (2024). *Metacognitive strategies, AI-based writing self-efficacy and writing anxiety in AI-assisted writing contexts: A structural equation modeling analysis*.

Silva Falcón, D. S. (2023). *Automated tools analysis in the written production.* Universidad Técnica de Ambato-Facultad de Ciencias Humanas y de la Educación …,

Tangpermpoon, T. (2008). Integrated approaches to improve students writing skills for English major students. *ABAC Journal, 28*(2).

Tinio, V. L. (2002). *Survey of information and communication technology utilization in Philippine public high schools.* Retrieved on April, 15, 2008.

Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing, 16*(4), 255–272.

Wiyaka, W., Mujiyanto, J., & Rukmini, D. (2018). Students' perception on the usefulness of ICT-based language program. *English Language Teaching, 11*(2), 53–60.

Wu, W., & Wu, P. (2008). Creating an authentic EFL learning environment to enhance student motivation to study English. *Asian EFL Journal, 10*(4), 211–226.

Wu, X., Yang, H., Liu, J., & Liu, Z. (2022). English use anxiety, motivation, self-efficacy, and their predictive effects on Chinese top university students' English achievements. *Frontiers in Psychology, 13*, 953600.

Xu, Z., Banerjee, M., Ramirez, G., Zhu, G., & Wijekumar, K. (2019). The effectiveness of educational technology applications on adult English language learners' writing quality: A meta-analysis. *Computer Assisted Language Learning, 32*(1–2), 132–162.

Yundayani, A., Susilawati, S., & Chairunnisa, C. (2019). Investigating the effect of Canva on students' writing skills. *English Review: Journal of English Education, 7*(2), 169–176.

Yunus, M. M., Nordin, N., Salehi, H., Embi, M. A., & Salehi, Z. (2013). The use of information and communication technology (ICT) in teaching ESL writing skills. *English Language Teaching, 6*(7), 1–8.

Yusuf, M. (2011). The impact of self-efficacy, achievement motivation, and self-regulated learning strategies on students' academic achievement. *Procedia-Social and Behavioral Sciences, 15*, 2623–2626.

Zaini, A., & Mazdayasna, G. (2015). The impact of computer-based instruction on the development of EFL learners' writing skills. *Journal of Computer Assisted Learning, 31*(6), 516–528.

Zamel, V. (1982). Writing: The process of discovering meaning. *TESOL Quarterly, 16*(2), 195–209.

Zhang, Z.-J., Zhang, C.-L., Zhang, X.-G., Liu, X.-M., Zhang, H., Wang, J., & Liu, S. (2015). Relationship between self-efficacy beliefs and achievement motivation in student nurses. *Chinese Nursing Research, 2*(2-3), 67–70.

Zhenpeng, Y. (2024). Application of artificial intelligence in computer network technology in the age of big data. *Journal of Artificial Intelligence Practice, 7*(1), 11–16.

Zrekat, Y., Abu Bakar, N., & Latif, H. (2016). The level of anxiety among Jordanian EFL undergraduates in oral communication performance. *Arab World English Journal (AWEJ), 7*(3), 188–202.

Appendix A

Cambridge English: Preliminary for Schools Writing Examiners use the following assessment scale, extracted from the one on the next page:

B1	Content	Communicative Achievement	Organisation	Language
5	All content is relevant to the task. Target reader is fully informed.	Uses the conventions of the communicative task to hold the target reader's attention and communicate straightforward ideas.	Text is generally well organised and coherent, using a variety of linking words and cohesive devices.	Uses a range of everyday vocabulary appropriately, with occasional inappropriate use of less common lexis. Uses a range of simple and some complex grammatical forms with a good degree of control. Errors do not impede communication.
4	Performance shares features of Bands 3 and 5.			
3	Minor irrelevances and/or omissions may be present. Target reader is on the whole informed.	Uses the conventions of the communicative task in generally appropriate ways to communicate straightforward ideas.	Text is connected and coherent, using basic linking words and a limited number of cohesive devices.	Uses everyday vocabulary generally appropriately, while occasionally overusing certain lexis. Uses simple grammatical forms with a good degree of control. While errors are noticeable, meaning can still be determined.
2	Performance shares features of Bands 1 and 3.			
1	Irrelevances and misinterpretation of task may be present. Target reader is minimally informed.	Produces text that communicates simple ideas in simple ways.	Text is connected using basic, high-frequency linking words.	Uses basic vocabulary reasonably appropriately. Uses simple grammatical forms with some degree of control. Errors may impede meaning at times.
0	Content is totally irrelevant. Target reader is not informed.	Performance below Band 1		

Appendix B

Essay Rating Scale: FL Composition Profile

	Score	Criteria
Content	27–30	*Excellent to very good:* knowledgeable; substantive, thorough development of thesis; relevant to topic assigned.
	22–26	Good to average: some knowledge of subject; adequate range; limited thematic development; mostly relevant to topic, but lacks detail.
	17–21	*Fair to poor:* limited knowledge of subject; minimal substance; poor thematic development.
	13–16	*Very poor:* shows little or no knowledge of subject; inadequate quantity; not relevant, or not enough to rate.
Organization	18–20	*Excellent to very good:* fluent expression; clear statement of ideas; solid support; clear organization; logical and cohesive sequencing.
	14–17	*Good to average:* adequate fluency; main ideas clear but loosely organized; supporting material limited; sequencing logical but incomplete.
	10–13	*Fair to poor:* low fluency; ideas not well connected; logical sequencing and development lacking.
	7–9	*Very poor:* ideas not communicated; organization lacking, or not enough to rate.
Grammar	22–25	*Excellent to very good:* accurate use of relatively complex structures; few errors in agreement, number, tense, word order, articles, pronouns, prepositions.
	18–21	*Good to average:* simple constructions used effectively; some problems in use of complex constructions; errors in agreement, number, tense, word order, articles, pronouns, prepositions.
	11–17	*Fair to poor:* significant defects in use of complex constructions; frequent errors in agreement, number, tense, negation, word order, articles, pronouns, prepositions; fragments and deletions; lack of accuracy interferes with meaning.
	5–10	*Very poor:* no mastery of simple sentence construction; text dominated by errors; does not communicate, or not enough to rate.
Vocabulary	18–20	*Excellent to very good:* complex range; accurate word/idiom choice; mastery of word forms; appropriate register.
	14–17	*Good to average:* adequate range; errors of word/idiom choice; effective transmission of meaning.
	10–13	*Fair to poor:* limited range; frequent word/idiom errors; inappropriate choice, usage; meaning not effectively communicated.
	7–9	*Very poor:* translation-based errors; little knowledge of target language vocabulary, or not enough to rate.
Mechanics	5	*Excellent to very good:* masters conventions of spelling, punctuation, capitalization, paragraph indentation, etc.
	4	*Good to average:* occasional errors in spelling, punctuation, capitalization, paragraph indentation, etc., which do not interfere with meaning.
	3	*Fair to poor:* frequent spelling, punctuation, capitalization, paragraphing errors; meaning disrupted by formal problems.
	2	*Very poor:* no mastery of conventions due to frequency of mechanical errors, or not enough to rate.
Total	_____/100	

Note. This scale was adapted from *Composición, Proceso y Síntesis* by G. Valdes and T. Dvorak, 1989, New York: McGraw-Hill. Copyright 1989 by McGraw-Hill, Inc. Reproduced by permission. (The original version appears in Jacobs, Zinkgraf, Wormuth, Hartfiel, and Hughey, 1981.)

Appendix C

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
1.1. I can use Hemingway to complete writing assignments.					
1.2. I can use appropriate vocabulary to express opinions when writing.					
1.3. I can continue the assignment after dealing with difficulties.					
1.4. I can generate a lot of ideas for writing assignments.					
1.5. I can write complete sentences without grammatical mistakes.					
2.1. Hemingway App gives me chances to practice writing.					
2.2. Hemingway App allows me to familiarize myself with the topic before writing.					
2.3. I feel under the pressure to complete my work within a time limit.					
2.4. I do not feel under the pressure to create the first perfect draft.					
2.5. I do not feel stressed when receiving feedback from my peers and teacher.					
2.6. I do not have difficulties using Hemingway to present my writing.					
3.1. The workload for writing assignments using the Hemingway App is reasonable.					
3.2. It is convenient to save work on Hemingway App.					
3.3. Hemingway App makes me more responsible for my own writing.					
3.4. Hemingway App helps me track my writing progress.					
3.5. Hemingway App makes the writing process more interesting.					
3.6. Hemingway App helps me realize my strengths and weaknesses.					
4.1. Hemingway App is an effective supplementary tool for writing.					
4.2. I would like to use Hemingway App to learn writing skills.					